PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

D4

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 9/02, 9/04, 9/00, 9/24, 9/10, 9/88, 9/08, 15/82

A2

(11) International Publication Number: **WO 98/11205**

(43) International Publication Date: 19 March 1998 (19.03.98)

(21) International Application Number: PCT/NZ97/00112

(22) International Filing Date: 10 September 1997 (10.09.97)

(30) Priority Data:
08/713,000 11 September 1996 (11.09.96) US

(71) Applicants: GENESIS RESEARCH & DEVELOPMENT CORPORATION LIMITED [NZ/NZ]; 1 Fox Street, Parnell, Auckland (NZ). FLETCHER CHALLENGE FORESTS LIMITED [NZ/NZ]; 585 Great South Road, Penrose, Auckland (NZ).

(72) Inventors: BLOKSBERG, Leonard, Nathan; 5A Korau Road, Greenlane, Auckland (NZ). GRIERSON, Alistair, Wallace; 1/24 Medina Place, Bucklands Beach, Auckland (NZ). HAVUKKALA, Ilkka, Jaakko; 3/121 Atkin Avenue, Mission Bay, Auckland (NZ).

(74) Agents: BENNETT, Michael, Roy et al.; Russell McVeagh West-Walker, The Todd Building, Level 5, 171-177 Lambton Quay, Wellington 6001 (NZ).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MATERIALS AND METHODS FOR THE MODIFICATION OF PLANT LIGNIN CONTENT

(57) Abstract

Novel isolated DNA sequences associated with the lignin biosynthetic pathway are provided, together with DNA constructs including such sequences. Methods for the modulation of lignin content in plants are also disclosed, the methods comprising incorporating one or more of the inventive DNA sequences or a sequence complementary to an inventive DNA sequence into the genome of a plant.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MATERIALS AND METHODS FOR THE MODIFICATION OF PLANT LIGNIN CONTENT

Technical Field of the Invention

This invention relates to the field of modification of lignin content and composition in plants. More particularly, this invention relates to enzymes involved in the lignin biosynthetic pathway and nucleotide sequences encoding such enzymes.

Background of the Invention

Lignin is an insoluble polymer which is primarily responsible for the rigidity of plant stems. Specifically, lignin serves as a matrix around the polysaccharide components of some plant cell walls. The higher the lignin content, the more rigid the plant. For example, tree species synthesize large quantities of lignin, with lignin constituting between 20% to 30% of the dry weight of wood. In addition to providing rigidity, lignin aids in water transport within plants by rendering cell walls hydrophobic and water impermeable. Lignin also plays a role in disease resistance of plants by impeding the penetration and propagation of pathogenic agents.

The high concentration of lignin in trees presents a significant problem in the paper industry wherein considerable resources must be employed to separate lignin from the cellulose fiber needed for the production of paper. Methods typically employed for the removal of lignin are highly energy- and chemical-intensive, resulting in increased costs and increased levels of undesirable waste products. In the U.S. alone, about 20 million tons of lignin are removed from wood per year.

Lignin is largely responsible for the digestibility, or lack thereof, of forage crops, with small increases in plant lignin content resulting in relatively high decreases in digestibility. For example, crops with reduced lignin content provide more efficient forage for cattle, with the yield of milk and meat being higher relative to the amount of forage crop consumed. During normal plant growth, the increase in dry matter content is accompanied by a corresponding decrease in digestibility. When deciding on the optimum time to harvest forage crops, farmers must therefore chose between a high yield of less digestible material and a lower yield of more digestible material.

For some applications, an increase in lignin content is desirable since increasing the lignin content of a plant would lead to increased mechanical strength of wood, changes in its color and increased resistance to rot. Mycorrhizal species composition and abundance may also be favorably manipulated by modifying lignin content and structural composition.

As discussed in detail below, lignin is formed by polymerization of at least three different monolignols which are synthesized in a multistep pathway, each step in the pathway being catalyzed by a different enzyme. It has been shown that manipulation of the number of copies of genes encoding certain enzymes, such as cinnamyl alcohol dehydrogenase (CAD) and caffeic acid 3-O-methyltransferase (COMT) results in modification of the amount of lignin produced; see, for example, U.S. Patent No. 5,451,514 and PCT publication no. WO 94/23044. Furthermore, it has been shown that antisense expression of sequences encoding CAD in poplar leads to the production of lignin having a modified composition (Grand, C. et al. Planta (Berl.) 163:232-237 (1985)).

While DNA sequences encoding some of the enzymes involved in the lignin biosynthetic pathway have been isolated for certain species of plants, genes encoding many of the enzymes in a wide range of plant species have not yet been identified. Thus there remains a need in the art for materials useful in the modification of lignin content and composition in plants and for methods for their use.

Summary of the Invention

Briefly, the present invention provides isolated DNA sequences obtainable from eucalyptus and pine which encode enzymes involved in the lignin biosynthetic pathway, DNA constructs including such sequences, and methods for the use of such constructs. Transgenic plants having altered lignin content and composition are also provided.

In a first aspect, the present invention provides isolated DNA sequences coding for the following enzymes isolated from eucalyptus and pine: cinnamate 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), phenolase (PNL), O-methyl transferase (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL), coniferol

glucosyl transferase (CGT), coniferin *beta*-glucosidase (CBG), laccase (LAC) and peroxidase (POX), together with ferulate-5-hydroxylase (F5H) from eucalyptus. In one embodiment, the isolated DNA sequences comprise a nucleotide sequence selected from the group consisting of: (a) sequences recited in SEQ ID NO: 3, 13, 16-70, and 72-88; (b) complements of the sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88; (c) reverse complements of the sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88; (d) reverse sequences of the sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88; and (e) sequences having at least about a 99% probability of being the same as a sequence of (a) – (d) as measured by the computer algorithm FASTA.

In another aspect, the invention provides DNA constructs comprising a DNA sequence of the present invention, either alone, in combination with one or more of the inventive sequences or in combination with one or more known DNA sequences; together with transgenic cells comprising such constructs.

In a related aspect, the present invention provides DNA constructs comprising, in the 5'-3' direction, a gene promoter sequence; an open reading frame coding for at least a functional portion of an enzyme encoded by the inventive DNA sequences or variants thereof; and a gene termination sequence. The open reading frame may be orientated in either a sense or antisense direction. DNA constructs comprising a non-coding region of a gene coding for an enzyme encoded by the above DNA sequences or a nucleotide sequence complementary to a non-coding region, together with a gene promoter sequence and a gene termination sequence, are also provided. Preferably, the gene promoter and termination sequences are functional in a host plant. Most preferably, the gene promoter and termination sequences are those of the original enzyme genes but others generally used in the art, such as the Cauliflower Mosaic Virus (CMV) promoter, with or without enhancers, such as the Kozak sequence or Omega enhancer, and *Agrobacterium tumefaciens* nopalin synthase terminator may be usefully employed in the present invention. Tissue-specific promoters may be employed in order to target expression to one or more desired tissues. In a preferred embodiment, the gene promoter sequence provides for transcription in xylem. The DNA construct may further include a marker for the identification of transformed cells.

In a further aspect, transgenic plant cells comprising the DNA constructs of the present invention are provided, together with plants comprising such transgenic cells, and fruits and seeds of such plants.

In yet another aspect, methods for modulating the lignin content and composition of a plant are provided, such methods including stably incorporating into the genome of the plant a DNA construct of the present invention. In a preferred embodiment, the target plant is a woody plant, preferably selected from the group consisting of eucalyptus and pine species, most preferably from the group consisting of *Eucalyptus grandis* and *Pinus radiata*. In a related aspect, a method for producing a plant having altered lignin content is provided, the method comprising transforming a plant cell with a DNA construct of the present invention to provide a transgenic cell, and cultivating the transgenic cell under conditions conducive to regeneration and mature plant growth.

In yet a further aspect, the present invention provides methods for modifying the activity of an enzyme in a plant, comprising stably incorporating into the genome of the plant a DNA construct of the present invention. In a preferred embodiment, the target plant is a woody plant, preferably selected from the group consisting of eucalyptus and pine species, most preferably from the group consisting of *Eucalyptus grandis* and *Pinus radiata*.

The above-mentioned and additional features of the present invention and the manner of obtaining them will become apparent, and the invention will be best understood by reference to the following more detailed description, read in conjunction with the accompanying drawing.

Brief Description of the Figures

Fig. 1 is a schematic overview of the lignin biosynthetic pathway.

Detailed Description

Lignin is formed by polymerization of at least three different monolignols, primarily *para*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. While these three types of lignin subunits are well known, it is possible that slightly different variants of these subunits may be involved in the lignin biosynthetic pathway in various

plants. The relative concentration of these residues in lignin varies between different plant species and within species. In addition, the composition of lignin may also vary between different tissues within a specific plant. The three monolignols are derived from phenylalanine in a multistep process and are believed to be polymerized into lignin by a free radical mechanism.

Fig. 1 shows the different steps in the biosynthetic pathway for coniferyl alcohol together with the enzymes responsible for catalyzing each step. *para*-Coumaryl alcohol and sinapyl alcohol are synthesized by similar pathways. Phenylalanine is first deaminated by phenylalanine ammonia-lyase (PAL) to give cinnamate which is then hydroxylated by cinnamate 4-hydroxylase (C4H) to form *p*-coumarate. *p*-Coumarate is hydroxylated by coumarate 3-hydroxylase to give caffeate. The newly added hydroxyl group is then methylated by O-methyl transferase (OMT) to give ferulate which is conjugated to coenzyme A by 4-coumarate:CoA ligase (4CL) to form feruloyl-CoA. Reduction of feruloyl-CoA to coniferaldehyde is catalyzed by cinnamoyl-CoA reductase (CCR). Coniferaldehyde is further reduced by the action of cinnamyl alcohol dehydrogenase (CAD) to give coniferyl alcohol which is then converted into its glucosylated form for export from the cytoplasm to the cell wall by coniferol glucosyl transferase (CGT). Following export, the de-glucosylated form of coniferyl alcohol is obtained by the action of coniferin *beta*-glucosidase (CBG). Finally, polymerization of the three monolignols to provide lignin is catalyzed by phenolase (PNL), laccase (LAC) and peroxidase (POX).

The formation of sinapyl alcohol involves an additional enzyme, ferulate-5-hydroxylase (F5H). For a more detailed review of the lignin biosynthetic pathway, see: Whetton, R. and Sederoff, R., The Plant Cell, 7:1001-1013 (1995).

Quantitative and qualitative modifications in plant lignin content are known to be induced by external factors such as light stimulation, low calcium levels and mechanical stress. Synthesis of new types of lignins, sometimes in tissues not normally lignified, can also be induced by infection with pathogens. In addition to lignin, several other classes of plant products are derived from phenylalanine, including flavonoids, coumarins, stilbenes and benzoic acid derivatives, with the initial steps in the synthesis of all these compounds being the same. Thus modification of the action of PAL, C4H and 4CL may affect the synthesis of other plant products in addition to lignin.

Using the methods and materials of the present invention, the lignin content of a plant can be increased by incorporating additional copies of genes encoding enzymes involved in the lignin biosynthetic pathway into the genome of the target plant. Similarly, a decrease in lignin content can be obtained by transforming the target plant with antisense copies of such genes. In addition, the number of copies of genes encoding for different enzymes in the lignin biosynthetic pathway can be manipulated to modify the relative amount of each monolignol synthesized, thereby leading to the formation of lignin having altered composition. The alteration of lignin composition would be advantageous, for example, in tree processing for paper, and may also be effective in altering the palatability of wood materials to rotting fungi.

In one embodiment, the present invention provides isolated complete or partial DNA sequences encoding, or partially encoding, enzymes involved in the lignin biosynthetic pathway, the DNA sequences being obtainable from eucalyptus and pine. Specifically, the present invention provides isolated DNA sequences encoding the enzymes CAD (SEQ ID NO: 1, 30), PAL (SEQ ID NO: 16), C4H (SEQ ID NO: 17), C3H (SEQ ID NO: 18), F5H (SEQ ID NO: 19-21), OMT (SEQ ID NO: 22-25), CCR (SEQ ID NO: 26-29), CGT (SEQ ID NO: 31-33), CBG (SEQ ID NO: 34), PNL (SEQ ID NO: 35, 36), LAC (SEQ ID NO: 37-41) and POX (SEQ ID NO: 42-44) from *Eucalyptus grandis*; and the enzymes C4H (SEQ ID NO: 2, 3, 48, 49), C3H (SEQ ID NO: 4, 50-52), PNL (SEQ ID NO: 5, 81), OMT (SEQ ID NO: 6, 53-55), CAD (SEQ ID NO: 7, 71), CCR (SEQ ID NO: 8, 58-70), PAL (SEQ ID NO: 9-11,45-47), 4CL (SEQ ID NO: 12, 56, 57), CGT (SEQ ID NO: 72), CBG (SEQ ID NO: 73-80), LAC (SEQ ID NO: 82-84) and POX (SEQ ID NO: 13, 85-88) from *Pinus radiata*. Complements of such isolated DNA sequences, reverse complements of such isolated DNA sequences and reverse sequences of such isolated DNA sequences, together with variants of such sequences, are also provided. DNA sequences encompassed by the present invention include cDNA, genomic DNA, recombinant DNA and wholly or partially chemically synthesized DNA molecules.

The definition of the terms "complement", "reverse complement" and "reverse sequence", as used herein, is best illustrated by the following example. For the sequence 5' AGGACC 3', the complement, reverse complement and reverse sequence are as follows:

complement	3' TCCTGG 5'
reverse complement	3' GGTCCT 5'
reverse sequence	5' CCAGGA 3'.

As used herein, the term "variant" covers any sequence which exhibits at least about 50%, more preferably at least about 70% and, more preferably yet, at least about 90% identity to a sequence of the present invention. Most preferably, a "variant" is any sequence which has at least about a 99% probability of being the same as the inventive sequence. The probability for DNA sequences is measured by the computer algorithm FASTA (version 2.0u4, February 1996; Pearson W. R. et al., Proc. Natl. Acad. Sci., 85:2444-2448, 1988), the probability for translated DNA sequences is measured by the computer algorithm TBLASTX and that for protein sequences is measured by the computer algorithm BLASTP (Altschul, S. F. et al. J. Mol. Biol., 215:403-410, 1990). The term "variants" thus encompasses sequences wherein the probability of finding a match by chance (smallest sum probability) in a database, is less than about 1% as measured by any of the above tests.

Variants of the isolated sequences from other eucalyptus and pine species, as well as from other commercially important species utilized by the lumber industry, are contemplated. These include the following gymnosperms, by way of example: loblolly pine *Pinus taeda*, slash pine *Pinus elliotti*, sand pine *Pinus clausa*, longleaf pine *Pinus palustrus*, shortleaf pine *Pinus echinata*, ponderosa pine *Pinus ponderosa*, Jeffrey pine *Pinus jeffrey*, red pine *Pinus resinosa*, pitch pine *Pinus rigida*, jack pine *Pinus banksiana*, pond pine *Pinus serotina*, Eastern white pine *Pinus strobus*, Western white pine *Pinus monticola*, sugar pine *Pinus lambertiana*, Virginia pine *Pinus virginiana*, lodgepole pine *Pinus contorta*, Caribbean pine *Pinus caribaea*, *P. pinaster*, Calabrian pine *P. brutia*, Afghan pine *P. eldarica*, Coulter pine *P. coulteri*, European pine *P. nigra* and *P. sylvestris*; Douglas-fir *Pseudotsuga menziesii*; the hemlocks which include Western hemlock *Tsuga heterophylla*, Eastern hemlock *Tsuga canadensis*, Mountain hemlock *Tsuga mertensiana*; the spruces which include the Norway spruce *Picea abies*, red spruce *Picea rubens*, white spruce *Picea glauca*, black spruce *Picea mariana*, Sitka spruce *Picea sitchensis*, Englemann spruce *Picea engelmanni*, and blue spruce *Picea pungens*; redwood *Sequoia sempervirens*; the true firs include the Alpine fir *Abies lasiocarpa*, silver fir *Abies amabilis*, grand fir *Abies grandis*, noble fir *Abies procera*, white fir *Abies concolor*, California red fir *Abies magnifica*, and balsam fir *Abies balsamea*, the cedars which include the Western red cedar *Thuja plicata*, incense

cedar *libocedrus decurrens*, Northern white cedar *Thuja occidentalis*, Port Orford cedar *Chamaecyparis lawsoniona*, Atlantic white cedar *Chamaecyparis thyoides*, Alaska yellow-cedar *Chamaecyparis nootkatensis*, and Eastern red cedar *Huniperus virginiana*; the larches which include Eastern larch *Larix laricina*, Western larch *Larix occidentalis*, European larch *Larix decidua*, Japanese larch *Larix leptolepis*, and Siberian larch *Larix siberica*; bold cypress *Taxodium distichum* and Giant sequoia *Sequoia gigantea*;

and the following angiosperms, by way of example:

Eucalyptus alba, *E. bancroftii*, *E. botyroides*, *E. bridgesiana*, *E. calophylla*, *E. camaldulensis*, *E. citriodora*, *E. cladocalyx*, *E. coccifera*, *E. curtisii*, *E. dalrympleana*, *E. deglupta*, *E. delagatensis*, *E. diversicolor*, *E. dunnii*, *E. ficifolia*, *E. globulus*, *E. gomphocephala*, *E. gunnii*, *E. henryi*, *E. laevopinea*, *E. macarthurii*, *E. macrorhyncha*, *E. maculata*, *E. marginata*, *E. megacarpa*, *E. melliodora*, *E. nicholii*, *E. nitens*, *E. nova-anglica*, *E. obliqua*, *E. obtusiflora*, *E. oreades*, *E. pauciflora*, *E. polybractea*, *E. regnans*, *E. resinifera*, *E. robusta*, *E. rudis*, *E. saligna*, *E. sideroxylon*, *E. stuartiana*, *E. tereticornis*, *E. torelliana*, *E. urnigera*, *E. urophylla*, *E. viminalis*, *E. viridis*, *E. wandoo* and *E. youmanni*.

The inventive DNA sequences may be isolated by high throughput sequencing of cDNA libraries such as those prepared from *Eucalyptus grandis* and *Pinus radiata* as described below in Examples 1 and 2. Alternatively, oligonucleotide probes based on the sequences provided in SEQ ID NO: 1-13 and 16-88 can be synthesized and used to identify positive clones in either cDNA or genomic DNA libraries from *Eucalyptus grandis* and *Pinus radiata*, or from other gymnosperms and angiosperms including those identified above, by means of hybridization or PCR techniques. Probes can be shorter than the sequences provided herein but should be at least about 10, preferably at least about 15 and most preferably at least about 20 nucleotides in length. Hybridization and PCR techniques suitable for use with such oligonucleotide probes are well known in the art. Positive clones may be analyzed by restriction enzyme digestion, DNA sequencing or the like.

In addition, the DNA sequences of the present invention may be generated by synthetic means using techniques well known in the art. Equipment for automated synthesis of oligonucleotides is commercially available from suppliers such as Perkin Elmer/Applied Biosystems Division (Foster City, CA) and may be operated according to the manufacturer's instructions.

In one embodiment, the DNA constructs of the present invention include an open reading frame coding for at least a functional portion of an enzyme encoded by a nucleotide sequence of the present invention or a variant thereof. As used herein, the "functional portion" of an enzyme is that portion which contains the active site essential for affecting the metabolic step, *i.e.* the portion of the molecule that is capable of binding one or more reactants or is capable of improving or regulating the rate of reaction. The active site may be made up of separate portions present on one or more polypeptide chains and will generally exhibit high substrate specificity. The term "enzyme encoded by a nucleotide sequence" as used herein, includes enzymes encoded by a nucleotide sequence which includes the partial isolated DNA sequences of the present invention.

For applications where amplification of lignin synthesis is desired, the open reading frame is inserted in the DNA construct in a sense orientation, such that transformation of a target plant with the DNA construct will lead to an increase in the number of copies of the gene and therefore an increase in the amount of enzyme. When down-regulation of lignin synthesis is desired, the open reading frame is inserted in the DNA construct in an antisense orientation, such that the RNA produced by transcription of the DNA sequence is complementary to the endogenous mRNA sequence. This, in turn, will result in a decrease in the number of copies of the gene and therefore a decrease in the amount of enzyme. Alternatively, regulation can be achieved by inserting appropriate sequences or subsequences (e.g. DNA or RNA) in ribozyme constructs.

In a second embodiment, the inventive DNA constructs comprise a nucleotide sequence including a non-coding region of a gene coding for an enzyme encoded by a DNA sequence of the present invention, or a nucleotide sequence complementary to such a non-coding region. As used herein the term "non-coding region" includes both transcribed sequences which are not translated, and non-transcribed sequences within about 2000 base pairs 5' or 3' of the translated sequences or open reading frames. Examples of non-coding regions which may be usefully employed in the inventive constructs include introns and 5'-non-coding leader sequences. Transformation of a target plant with such a DNA construct may lead to a reduction in the amount of lignin synthesized by the plant by the process of cosuppression, in a manner similar to that

discussed, for example, by Napoli et al. (Plant Cell 2:279-290, 1990) and de Carvalho Niebel et al. (Plant Cell 7:347-358, 1995).

The DNA constructs of the present invention further comprise a gene promoter sequence and a gene termination sequence, operably linked to the DNA sequence to be transcribed, which control expression of the gene. The gene promoter sequence is generally positioned at the 5' end of the DNA sequence to be transcribed, and is employed to initiate transcription of the DNA sequence. Gene promoter sequences are generally found in the 5' non-coding region of a gene but they may exist in introns (Luehrsen, K. R., Mol. Gen. Genet. 225:81-93, 1991) or in the coding region, as for example in PAL of tomato (Bloksberg, 1991, Studies on the Biology of Phenylalanine Ammonia Lyase and Plant Pathogen Interaction. Ph.D. Thesis, Univ. of California, Davis, University Microfilms International order number 9217564). When the construct includes an open reading frame in a sense orientation, the gene promoter sequence also initiates translation of the open reading frame. For DNA constructs comprising either an open reading frame in an antisense orientation or a non-coding region, the gene promoter sequence consists only of a transcription initiation site having a RNA polymerase binding site.

A variety of gene promoter sequences which may be usefully employed in the DNA constructs of the present invention are well known in the art. The promoter gene sequence, and also the gene termination sequence, may be endogenous to the target plant host or may be exogenous, provided the promoter is functional in the target host. For example, the promoter and termination sequences may be from other plant species, plant viruses, bacterial plasmids and the like. Preferably, gene promoter and termination sequences are from the inventive sequences themselves.

Factors influencing the choice of promoter include the desired tissue specificity of the construct, and the timing of transcription and translation. For example, constitutive promoters, such as the 35S Cauliflower Mosaic Virus (CaMV 35S) promoter, will affect the activity of the enzyme in all parts of the plant. Use of a tissue specific promoter will result in production of the desired sense or antisense RNA only in the tissue of interest. With DNA constructs employing inducible gene promoter sequences, the rate of RNA polymerase binding and initiation can be modulated by external stimuli, such as light, heat, anaerobic stress, alteration in nutrient conditions

and the like. Temporally regulated promoters can be employed to effect modulation of the rate of RNA polymerase binding and initiation at a specific time during development of a transformed cell. Preferably, the original promoters from the enzyme gene in question, or promoters from a specific tissue-targeted gene in the organism to be transformed, such as eucalyptus or pine are used. Other examples of gene promoters which may be usefully employed in the present invention include, mannopine synthase (mas), octopine synthase (ocs) and those reviewed by Chua et al. (Science, 244:174-181, 1989).

The gene termination sequence, which is located 3' to the DNA sequence to be transcribed, may come from the same gene as the gene promoter sequence or may be from a different gene. Many gene termination sequences known in the art may be usefully employed in the present invention, such as the 3' end of the *Agrobacterium tumefaciens* nopaline synthase gene. However, preferred gene terminator sequences are those from the original enzyme gene or from the target species to be transformed.

The DNA constructs of the present invention may also contain a selection marker that is effective in plant cells, to allow for the detection of transformed cells containing the inventive construct. Such markers, which are well known in the art, typically confer resistance to one or more toxins. One example of such a marker is the NPTII gene whose expression results in resistance to kanamycin or hygromycin, antibiotics which is usually toxic to plant cells at a moderate concentration (Rogers et al. in Methods for Plant Molecular Biology, A. Weissbach and H. Weissbach, eds., Academic Press Inc., San Diego, CA (1988)). Alternatively, the presence of the desired construct in transformed cells can be determined by means of other techniques well known in the art, such as Southern and Western blots.

Techniques for operatively linking the components of the inventive DNA constructs are well known in the art and include the use of synthetic linkers containing one or more restriction endonuclease sites as described, for example, by Maniatis et al., (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). The DNA construct of the present invention may be linked to a vector having at least one replication system, for example, *E. coli*, whereby after each manipulation, the resulting construct can be cloned and sequenced and the correctness of the manipulation determined.

The DNA constructs of the present invention may be used to transform a variety of plants, both monocotyledonous (e.g. grasses, corn, grains, oat, wheat and barley), dicotyledonous (e.g. *Arabidopsis*, tobacco, legumes, alfalfa, oaks, eucalyptus, maple), and Gymnosperms (e.g. Scots pine (Aronen, Finnish Forest Res. Papers, vol. 595, 1996), white spruce (Ellis et al., Biotechnology 11:94-92, 1993), larch (Huang et al., In Vitro Cell 27:201-207, 1991). In a preferred embodiment, the inventive DNA constructs are employed to transform woody plants, herein defined as a tree or shrub whose stem lives for a number of years and increases in diameter each year by the addition of woody tissue. Preferably the target plant is selected from the group consisting of eucalyptus and pine species, most preferably from the group consisting of *Eucalyptus grandis* and *Pinus radiata*. As discussed above, transformation of a plant with a DNA construct including an open reading frame coding for an enzyme encoded by an inventive DNA sequence wherein the open reading frame is orientated in a sense direction will lead to an increase in lignin content of the plant or, in some cases, to a decrease by cosuppression. Transformation of a plant with a DNA construct comprising an open reading frame in an antisense orientation or a non-coding (untranslated) region of a gene will lead to a decrease in the lignin content of the transformed plant.

Techniques for stably incorporating DNA constructs into the genome of target plants are well known in the art and include *Agrobacterium tumefaciens* mediated introduction, electroporation, protoplast fusion, injection into reproductive organs, injection into immature embryos, high velocity projectile introduction and the like. The choice of technique will depend upon the target plant to be transformed. For example, dicotyledonous plants and certain monocots and gymnosperms may be transformed by *Agrobacterium* Ti plasmid technology, as described, for exampie by Bevan (Nucl. Acid Res. 12:8711-8721, 1984). Targets for the introduction of the DNA constructs of the present invention include tissues, such as leaf tissue, disseminated cells, protoplasts, seeds, embryos, meristematic regions; cotyledons, hypocotyls, and the like. One preferred method for transforming eucalyptus and pine is a biolistic method using pollen (see, for example, Aronen 1996, Finnish Forest Res. Papers vol. 595, 53pp) or easily regenerable embryonic tissues. Other transformation techniques which may be usefully employed in the inventive methods include those taught by Ellis et al. (Plant

Cell Reports, 8:16-20, 1989), Wilson et al. (Plant Cell Reports 7:704-707, 1989) and Tautorus et al. (Theor. Appl. Genet. 78:531-536, 1989).

Once the cells are transformed, cells having the inventive DNA construct incorporated in their genome may be selected by means of a marker, such as the kanamycin resistance marker discussed above. Transgenic cells may then be cultured in an appropriate medium to regenerate whole plants, using techniques well known in the art. In the case of protoplasts, the cell wall is allowed to reform under appropriate osmotic conditions. In the case of seeds or embryos, an appropriate germination or callus initiation medium is employed. For explants, an appropriate regeneration medium is used. Regeneration of plants is well established for many species. For a review of regeneration of forest trees see Dunstan et al., Somatic embryogenesis in woody plants. In: Thorpe, T.A. ed., 1995: in vitro embryogenesis of plants. Vol. 20 in Current Plant Science and Biotechnology in Agriculture, Chapter 12, pp. 471-540. Specific protocols for the regeneration of spruce are discussed by Roberts et al., (Somatic Embryogenesis of Spruce. In: *Synseed. Applications of synthetic seed to crop improvement.* Redenbaugh, K., ed. CRC Press, Chapter 23, pp. 427-449, 1993). The resulting transformed plants may be reproduced sexually or asexually, using methods well known in the art, to give successive generations of transgenic plants.

As discussed above, the production of RNA in target plant cells can be controlled by choice of the promoter sequence, or by selecting the number of functional copies or the site of integration of the DNA sequences incorporated into the genome of the target plant host. A target plant may be transformed with more than one DNA construct of the present invention, thereby modulating the lignin biosynthetic pathway for the activity of more than one enzyme, affecting enzyme activity in more than one tissue or affecting enzyme activity at more than one expression time. Similarly, a DNA construct may be assembled containing more than one open reading frame coding for an enzyme encoded by a DNA sequence of the present invention or more than one non-coding region of a gene coding for such an enzyme. The DNA sequences of the present inventive may also be employed in combination with other known sequences encoding enzymes involved in the lignin biosynthetic pathway. In this manner, it may be possible to add a lignin biosynthetic pathway to a non-woody plant to produce a new woody plant.

The isolated DNA sequences of the present invention may also be employed as probes to isolate DNA sequences encoding enzymes involved in the lignin synthetic pathway from other plant species, using techniques well known to those of skill in the art.

The following examples are offered by way of illustration and not by way of limitation.

Example 1

Isolation and Characterization of cDNA Clones from *Eucalyptus grandis*

Two *Eucalyptus grandis* cDNA expression libraries (one from a mixture of various tissues from a single tree and one from leaves of a single tree) were constructed and screened as follows.

mRNA was extracted from the plant tissue using the protocol of Chang et al. (Plant Molecular Biology Reporter 11:113-116 (1993)) with minor modifications. Specifically, samples were dissolved in CPC-RNAXB (100 mM Tris-Cl, pH 8,0; 25 mM EDTA; 2.0 M NaCl; 2%CTAB; 2% PVP and 0.05% Spermidine*3 HCl)and extracted with Chloroform:isoamyl alcohol, 24:1. mRNA was precipitated with ethanol and the total RNA preparate was purified using a Poly(A) Quik mRNA Isolation Kit (Stratagene, La Jolla, CA). A cDNA expression library was constructed from the purified mRNA by reverse transcriptase synthesis followed by insertion of the resulting cDNA clones in Lambda ZAP using a ZAP Express cDNA Synthesis Kit (Stratagene), according to the manufacturer's protocol. The resulting cDNAs were packaged using a Gigapack II Packaging Extract (Stratagene) employing 1 µl of sample DNA from the 5 µl ligation mix. Mass excision of the library was done using XL1-Blue MRF' cells and XLOLR cells (Stratagene) with ExAssist helper phage (Stratagene). The excised phagemids were diluted with NZY broth (Gibco BRL, Gaithersburg, MD) and plated out onto LB-kanamycin agar plates containing X-gal and isopropylthio-beta-galactoside (IPTG).

Of the colonies plated and picked for DNA miniprep, 99% contained an insert suitable for sequencing. Positive colonies were cultured in NZY broth with kanamycin and cDNA was purified by means of alkaline lysis and polyethylene glycol (PEG) precipitation. Agarose gel at 1% was used to screen sequencing templates for

chromosomal contamination. Dye primer sequences were prepared using a Turbo Catalyst 800 machine (Perkin Elmer/Applied Biosystems, Foster City, CA) according to the manufacturer's protocol.

DNA sequence for positive clones was obtained using an Applied Biosystems Prism 377 sequencer. cDNA clones were sequenced first from both the 5' end and, in some cases, also from the 3' end. For some clones, internal sequence was obtained using subcloned fragments. Subcloning was performed using standard procedures of restriction mapping and subcloning to pBluescript II SK+ vector.

The determined cDNA sequence was compared to known sequences in the EMBL database (release 46, March 1996) using the FASTA algorithm of February 1996 (version 2.0u4) (available on the Internet at the ftp site ftp://ftp.virginia.edu/pub/fasta/). Multiple alignments of redundant sequences were used to build up reliable consensus sequences. Based on similarity to known sequences from other plant species, the isolated DNA sequence (SEQ ID NO: 1) was identified as encoding a CAD enzyme.

In further studies, using the procedure described above, cDNA sequences encoding the following *Eucalyptus grandis* enzymes were isolated: PAL (SEQ ID NO: 16); C4H (SEQ ID NO: 17); C3H (SEQ ID NO: 18); F5H (SEQ ID NO: 19-21); OMT (SEQ ID NO: 22-25); CCR (SEQ ID NO: 26-29); CAD (SEQ ID NO: 30); CGT (SEQ ID NO: 31-33); CBG (SEQ ID NO: 34); PNL (SEQ ID NO: 35, 36); LAC (SEQ ID NO: 37-41); and POX (SEQ ID NO: 42-44).

Example 2

Isolation and Characterization of cDNA Clones from *Pinus radiata*

a) Isolation of cDNA clones by high through-put screening

A *Pinus radiata* cDNA expression library was constructed from xylem and screened as described above in Example 1. DNA sequence for positive clones was obtained using forward and reverse primers on an Applied Biosystems Prism 377 sequencer and the determined sequences were compared to known sequences in the database as described above.

Based on similarity to known sequences from other plant species, the isolated DNA sequences were identified as encoding the enzymes C4H (SEQ ID NO: 2 and 3), C3H (SEQ ID NO: 4), PNL (SEQ ID NO: 5), OMT (SEQ ID NO: 6), CAD (SEQ ID NO: 7), CCR (SEQ ID NO: 8), PAL (SEQ ID NO: 9-11) and 4CL (SEQ ID NO: 12).

In further studies, using the procedure described above, additional cDNA clones encoding the following *Pinus radiata* enzymes were isolated: PAL (SEQ ID NO: 45-47); C4H (SEQ ID NO: 48, 49); C3H (SEQ ID NO: 50-52); OMT (SEQ ID NO: 53-55); 4CL (SEQ ID NO: 56, 57); CCR (SEQ ID NO: 58-70); CAD (SEQ ID NO: 71); CGT (SEQ ID NO: 72); CBG (SEQ ID NO: 73-80); PNL (SEQ ID NO: 81); LAC (SEQ ID NO: 82-84); and POX (SEQ ID NO: 85-88).

b) Isolation of cDNA clones by PCR

Two PCR probes, hereinafter referred to as LNB010 and LNB011 (SEQ ID NO: 14 and 15, respectively) were designed based on conserved domains in the following peroxidase sequences previously identified in other species: vanpox, hvupox6, taepox, hvupox1, osapox, ntopox2, ntopox1, lespox, pokpox, luspox, athpox, hrpox, spopox, and tvepox (Genbank accession nos. D11337, M83671, X56011, X58396, X66125, J02979, D11396, X71593, D11102, L07554, M58381, X57564, Z22920, and Z31011, respectively).

RNA was isolated from pine xylem and first strand cDNA was synthesized as described above. This cDNA was subjected to PCR using 4 μM LNB010, 4 μM LNB011, 1 x Kogen's buffer, 0.1 mg/ml BSA, 200 mM dNTP, 2 mM Mg^{2+}, and 0.1 U/μl of Taq polymerase (Gibco BRL). Conditions were 2 cycles of 2 min at 94 °C, 1 min at 55 °C and 1 min at 72 °C; 25 cycles of 1 min at 94 °C, 1 min at 55 °C, and 1 min at 72 °C; and 18 cycles of 1 min at 94 °C, 1 min at 55 °C, and 3 min at 72 °C in a Stratagene Robocycler. The gene was re-amplified in the same manner. A band of about 200 bp was purified from a TAE agarose gel using a Schleicher & Schuell Elu-Quik DNA purification kit and clones into a T-tailed pBluescript vector (Marchuk D. et al., Nucleic Acids Res. 19:1154, 1991). Based on similarity to known sequences, the isolated gene (SEQ ID NO: 13) was identified as encoding pine peroxidase (POX).

Example 3

Use of an O-methyltransferase (OMT) Gene to Modify Lignin Biosynthesis

a) Transformation of tobacco plants with a *Pinus radiata* OMT gene

Sense and anti-sense constructs containing a sequence including the coding region of OMT (SEQ ID NO: 53) from *Pinus radiata* were inserted into *Agrobacterium tumefaciens* LBA4301 (provided as a gift by Dr. C. Kado, University of California, Davis, CA) by direct transformation using published methods (see, An G, Ebert PR, Mitra A, Ha SB: Binary Vectors. In: Gelvin SB, Schilperoort RA (eds) Plant Molecular Biology Manual, Kluwer Academic Publishers, Dordrecht (1988)). The presence and integrity of the transgenic constructs were verified by restriction digestion and DNA sequencing.

Tobacco (*Nicotiana tabacum* cv. Samsun) leaf sections were transformed using the method of Horsch et al. (Science, 227:1229-1231, 1985). Five independent transformed plant lines were established for the sense construct and eight independent transformed plant lines were established for the anti-sense construct for OMT. Transformed plants containing the appropriate lignin gene construct were verified using Southern blot experiments. A "+" in the column labeled "Southern" in Table 1 below indicates that the transformed plant lines were confirmed as independent transformed lines.

b) Expression of *Pinus* OMT in transformed plants

Total RNA was isolated from each independent transformed plant line created with the OMT sense and anti-sense constructs. The RNA samples were analysed in Northern blot experiments to determine the level of expression of the transgene in each transformed line. The data shown in the column labeled "Northern" in Table 1 shows that the transformed plant lines containing the sense and anti-sense constructs for OMT all exhibited high levels of expression, relative to the background on the Northern blots. OMT expression in sense plant line number 2 was not measured because the RNA sample showed signs of degradation. There was no detectable hybridisation to RNA samples from empty vector-transformed control plants.

c) Modulation of OMT enzyme activity in transformed plants

The total activity of OMT enzyme, encoded by the *Pinus* OMT gene and by the endogenous tobacco OMT gene, in transformed tobacco plants was analysed for each transformed plant line created with the OMT sense and anti-sense constructs. Crude protein extracts were prepared from each transformed plant and assayed using the method of Zhang et al. (Plant Physiol., 113:65-74, 1997). The data contained in the column labeled "Enzyme" in Table 1 shows that the transformed plant lines containing the OMT sense construct generally had elevated OMT enzyme activity, with a maximum of 199%, whereas the transformed plant lines containing the OMT anti-sense construct generally had reduced OMT enzyme activity, with a minimum of 35%, relative to empty vector-transformed control plants. OMT enzyme activity was not estimated in sense plant line number 3.

d) Effects of *Pinus* OMT on lignin concentration in transformed plants

The concentration of lignin in the transformed tobacco plants was determined using the well-established procedure of thioglycolic acid extraction (see, Freudenberg et al. in "Constitution and Biosynthesis of Lignin", Springer-Verlag, Berlin, 1968). Briefly, whole tobacco plants, of an average age of 38 days, were frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle. 100 mg of frozen powder from one empty vector-transformed control plant line, the five independent transformed plant lines containing the sense construct for OMT and the eight independent transformed plant lines containing the anti-sense construct for OMT were extracted individually with methanol, followed by 10% thioglycolic acid and finally dissolved in 1 M NaOH. The final extracts were assayed for absorbance at 280 nm. The data shown in the column labelled "TGA" in Table 1 shows that the transformed plant lines containing the sense and the anti-sense OMT gene constructs all exhibited significantly decreased levels of lignin, relative to the empty vector-transformed control plant lines.

Table 1

plant line	transgene	orientation	Southern	Northern	Enzyme	TGA
1	control	na	+	blank	100	104
1	OMT	sense	+	2.9E+6	86	55
2	OMT	sense	+	na	162	58
3	OMT	sense	+	4.1E+6	na	63
4	OMT	sense	+	2.3E+6	142	66
5	OMT	sense	+	3.6E+5	199	75
1	OMT	anti-sense	+	1.6E+4	189	66
2	OMT	anti-sense	+	5.7E+3	35	70
3	OMT	anti-sense	+	8.0E+3	105	73
4	OMT	anti-sense	+	1.4E+4	109	74
5	OMT	anti-sense	+	2.5E+4	87	78
6	OMT	anti-sense	+	2.5E+4	58	84
7	OMT	anti-sense	+	2.5E+4	97	92
8	OMT	anti-sense	+	1.1E+4	151	94

These data clearly indicate that lignin concentration, as measured by the TGA assay, can be directly manipulated by either sense or anti-sense expression of a lignin biosynthetic gene such as OMT.

Example 4

Use of a 4-Coumarate:CoA ligase (4CL) Gene to Modify Lignin Biosynthesis

a) Transformation of tobacco plants with a *Pinus radiata* 4CL gene

Sense and anti-sense constructs containing a sequence including the coding region of 4CL (SEQ ID NO: 56) from *Pinus radiata* were inserted into *Agrobacterium tumefaciens* LBA4301 by direct transformation as described above. The presence and integrity of the transgenic constructs were verified by restriction digestion and DNA sequencing.

Tobacco (*Nicotiana tabacum* cv. Samsun) leaf sections were transformed as described above. Five independent transformed plant lines were established for the sense construct and eight independent transformed plant lines were established for the anti-sense construct for 4CL. Transformed plants containing the appropriate lignin gene construct were verified using Southern blot experiments. A "+" in the column

labeled "Southern" in Table 2 indicates that the transformed plant lines listed were confirmed as independent transformed lines.

b) Expression of *Pinus* 4CL in transformed plants

Total RNA was isolated from each independent transformed plant line created with the 4CL sense and anti-sense constructs. The RNA samples were analysed in Northern blot experiments to determine the level of expression of the transgene in each transformed line. The data shown in the column labelled "Northern" in Table 2 below shows that the transformed plant lines containing the sense and anti-sense constructs for 4CL all exhibit high levels of expression, relative to the background on the Northern blots. 4CL expression in anti-sense plant line number 1 was not measured because the RNA was not available at the time of the experiment. There was no detectable hybridisation to RNA samples from empty vector-transformed control plants.

c) Modulation of 4CL enzyme activity in transformed plants

The total activity of 4CL enzyme, encoded by the *Pinus* 4CL gene and by the endogenous tobacco 4CL gene, in transformed tobacco plants was analysed for each transformed plant line created with the 4CL sense and anti-sense constructs. Crude protein extracts were prepared from each transformed plant and assayed using the method of Zhang et al. (Plant Physiol., 113:65-74, 1997). The data contained in the column labeled "Enzyme" in Table 2 shows that the transformed plant lines containing the 4CL sense construct had elevated 4CL enzyme activity, with a maximum of 258%, and the transformed plant lines containing the 4CL anti-sense construct had reduced 4CL enzyme activity, with a minimum of 59%, relative to empty vector-transformed control plants.

d) Effects of *Pinus* 4CL on lignin concentration in transformed plants

The concentration of lignin in samples of transformed plant material was determined as described in Example 3. The data shown in the column labelled "TGA" in Table 2 shows that the transformed plant lines containing the sense and the anti-sense 4CL gene constructs all exhibited significantly decreased levels of lignin, relative to the empty vector-transformed control plant lines. These data clearly indicate that

lignin concentration, as measured by the TGA assay, can be directly manipulated by either sense or anti-sense expression of a lignin biosynthetic gene such as 4CL.

Table 2

plant line	transgene	orientation	Southern	Northern	Enzyme	TGA
1	control	na	+	blank	100	92
2	control	na	+	blank	100	104
1	4CL	sense	+	2.3E+4	169	64
2	4CL	sense	+	4.5E+4	258	73
3	4CL	sense	+	3.1E+4	174	77
4	4CL	sense	+	1.7E+4	164	80
5	4CL	sense	+	1.6E+4	184	92
1	4CL	anti-sense	+	na	59	75
2	4CL	anti-sense	+	1.0E+4	70	75
3	4CL	anti-sense	+	9.6E+3	81	80
4	4CL	anti-sense	+	1.2E+4	90	83
5	4CL	anti-sense	+	4.7E+3	101	88
6	4CL	anti-sense	+	3.9E+3	116	89
7	4CL	anti-sense	+	1.8E+3	125	94
8	4CL	anti-sense	+	1.7E+4	106	97

Example 5

Transformation of Tobacco using the Inventive Lignin Biosynthetic Genes

Sense and anti-sense constructs containing sequences including the coding regions of C3H (SEQ ID NO: 18), F5H (SEQ ID NO: 19), CCR (SEQ ID NO: 25) and CGT (SEQ ID NO: 31) from *Eucalyptus grandis*, and PAL (SEQ ID NO: 45 and 47), C4H (SEQ ID NO: 48 and 49), PNL (SEQ ID NO: 81) and LAC (SEQ ID NO: 83) from *Pinus radiata* were inserted into *Agrobacterium tumefaciens* LBA4301 by direct transformation as described above. The presence and integrity of the transgenic constructs were verified by restriction digestion and DNA sequencing.

Tobacco (*Nicotiana tabacum* cv. Samsun) leaf sections were transformed as described in Example 3. Up to twelve independent transformed plant lines were established for each sense construct and each anti-sense construct listed in the preceding paragraph. Transformed plants containing the appropriate lignin gene

construct were verified using Southern blot experiments. All of the transformed plant lines analysed were confirmed as independent transformed lines.

Example 6

Manipulation of Lignin Content in Transformed Plants

a) Determination of transgene expression by Northern blot experiments

Total RNA was isolated from each independent transformed plant line described in Example 5. The RNA samples were analysed in Northern blot experiments to determine the level of expression of the transgene in each transformed line. The column labelled "Northern" in Table 3 shows the level of transgene expression for all plant lines assayed, relative to the background on the Northern blots. There was no detectable hybridisation to RNA samples from empty vector-transformed control plants.

b) Determination of lignin concentration in transformed plants

The concentration of lignin in empty vector-transformed control plant lines and in up to twelve independent transformed lines for each sense construct and each anti-sense construct described in Example 5 was determined as described in Example 3. The column labelled "TGA" in Table 3 shows the thioglycolic acid extractable lignins for all plant lines assayed, expressed as the average percentage of TGA extractable lignins in transformed plants versus control plants. The range of variation is shown in parentheses.

Table 3

transgene	orientation	no. of lines	Northern	TGA
control	na	3	blank	100 (92-104)
C3H	sense	5	3.7E+4	74 (67-85)
F5H	sense	10	5.8E+4	70 (63-79)
F5H	anti-sense	9	5.8E+4	73 (35-93)
CCR	sense	1	na	74
CCR	anti-sense	2	na	74 (62-86)
PAL	sense	5	1.9E+5	77 (71-86)
PAL	anti-sense	4	1.5E+4	62 (37-77)
C4H	anti-sense	10	5.8E+4	86 (52-113)
PNL	anti-sense	6	1.2E+4	88 (70-114)
LAC	sense	5	1.7E+5	na
LAC	anti-sense	12	1.7E+5	88 (73-114)

Transformed plant lines containing the sense and the anti-sense lignin biosynthetic gene constructs all exhibited significantly decreased levels of lignin, relative to the empty vector-transformed control plant lines. The most dramatic effects on lignin concentration were seen in the F5H anti-sense plants with as little as 35% of the amount of lignin in control plants, and in the PAL anti-sense plants with as little as 37% of the amount of lignin in control plants. These data clearly indicate that lignin concentration, as measured by the TGA assay, can be directly manipulated by conventional anti-sense methodology and also by sense over-expression using the inventive lignin biosynthetic genes.

Example 7

Modulation of Lignin Enzyme Activity in Transformed Plants

The activities and substrate specificities of selected lignin biosynthetic enzymes were assayed in crude extracts from transformed tobacco plants containing sense and anti-sense constructs for PAL (SEQ ID NO: 45), PNL (SEQ ID NO: 81) and LAC (SEQ ID NO: 83) from *Pinus radiata*, and CGT (SEQ ID NO: 31) from *Eucalyptus grandis*.

Enzyme assays were performed using published methods for PAL (Southerton, S.G. and Deverall, B.J., Plant Path. 39:223-230, 1990), CGT (Vellekoop, P. et al., FEBS, 330:36-40, 1993), PNL (Espin, C.J. et al., Phytochemistry, 44:17-22, 1997) and

LAC (Bao, W. et al., Science, 260:672-674, 1993). The data shown in the column labelled "Enzyme" in Table 4 shows the average enzyme activity from replicate measures for all plant lines assayed, expressed as a percent of enzyme activity in empty vector-transformed control plants. The range of variation is shown in parentheses.

Table 4

transgene	orientation	no. of lines	Enzyme
control	na	3	100
PAL	sense	5	87 (60-124)
PAL	anti-sense	3	53 (38-80)
CGT	anti-sense	1	89
PNL	anti-sense	6	144 (41-279)
LAC	sense	5	78 (16-240)
LAC	anti-sense	11	64 (14-106)

All of the transformed plant lines, except the PNL anti-sense transformed plant lines, showed average lignin enzyme activities which were significantly lower than the activities observed in empty vector-transformed control plants. The most dramatic effects on lignin enzyme activities were seen in the PAL anti-sense transformed plant lines in which all of the lines showed reduced PAL activity and in the LAC anti-sense transformed plant lines which showed as little as 14% of the LAC activity in empty vector-transformed control plant lines.

Example 8

Functional Identification of Lignin Biosynthetic Genes

Sense constructs containing sequences including the coding regions for PAL (SEQ ID NO: 47), OMT (SEQ ID NO: 53), 4CL (SEQ ID NO: 56 and 57) and POX (SEQ ID NO: 86) from *Pinus radiata*, and OMT (SEQ ID NO: 23 and 24), CCR (SEQ ID NO: 26-28), CGT (SEQ ID NO: 31 and 33) and POX (SEQ ID NO: 42 and 44) from *Eucalyptus grandis* were inserted into the commercially available protein expression vector, pProEX-1 (Gibco BRL). The resultant constructs were transformed into *E. coli* XL1-Blue (Stratagene), which were then induced to produce recombinant protein by the addition of IPTG. Purified proteins were produced for the *Pinus* OMT and 4CL constructs and the *Eucalyptus* OMT and POX constructs using Ni column

chromatography (Janknecht, R. et al., Proc. Natl. Acad. Sci., 88:8972-8976, 1991). Enzyme assays for each of the purified proteins conclusively demonstrated the expected substrate specificity and enzymatic activity for the genes tested.

The data for two representative enzyme assay experiments, demonstrating the verification of the enzymatic activity of a *Pinus radiata* 4CL gene (SEQ ID NO: 56) and a *Pinus radiata* OMT gene (SEQ ID NO: 53), are shown in Table 5. For the 4CL enzyme, one unit equals the quantity of protein required to convert the substrate into product at the rate of 0.1 absorbance units per minute. For the OMT enzyme, one unit equals the quantity of protein required to convert 1 pmole of substrate to product per minute.

Table 5

transgene	purification step	total ml extract	total mg protein	total units activity	% yield activity	fold purification
4CL	crude	10 ml	51 mg	4200	100	1
	Ni column	4 ml	0.84 mg	3680	88	53
OMT	crude	10 ml	74 mg	4600	100	1
	Ni column	4 ml	1.2 mg	4487	98	60

The data shown in Table 5 indicate that both the purified 4CL enzyme and the purified OMT enzyme show high activity in enzyme assays, confirming the identification of the 4CL and OMT genes described in this application. Crude protein preparations from *E. coli* transformed with empty vector show no activity in either the 4CL or the OMT enzyme assay.

Although the present invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, changes and modifications can be carried out without departing from the scope of the invention which is intended to be limited only by the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Genesis Research and Development Corp. Ltd.

(ii) TITLE OF THE INVENTION: MATERIALS AND METHODS FOR THE MODIFICATION OF PLANT LIGNIN CONTENT

(iii) NUMBER OF SEQUENCES: 88

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Russell McVeagh West-Walker
(B) STREET: The Todd Building, Cnr Brandon Street & Lambton Quay
(C) CITY: Wellington
(D) STATE:
(E) COUNTRY: New Zealand
(F) ZIP:

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: Wordperfect 5.1

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Bennett, Michael Roy
(B) REGISTRATION NUMBER:
(C) REFERENCE/DOCKET NUMBER: 22315\MRB

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: +64 4 495 7740
(B) TELEFAX: +64 4 499 9306
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 535 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTTCGCGCTA CCGCATACTC CACCACCGCG TGCAGAAGAT GAGCTCGGAG GGTGGGAAGG 60
AGGATTGCCT CGGTTGGGCT GCCCGGGACC CTTCTGGGTT CCTCTCCCCN TACAAATTCA 120
CCCGCAGGCC GTGGGAAGCG AAGACGTCTC GATTAAGATC ACGCACTGTG GAGTGTGCTA 180
CGCAGATGTG GCTTGGACTA GGAATGTGCA GGGACACTCC AAGTATCCTC TGGTGCCGGG 240

GCACGAGATA GTTGGAATTG TGAAACAGGT TGGCTCCAGT GTCCAACGCT TCAAAGTTGG
300
CGATCATGTG GGGGTGGGAA CTTATGTCAA TTCATGCAGA GAGTGCGAGT ATTGCAATGA
360
CAGGCTAGAA GTCCAATGTG AAAAGTCGGT TATGACTTTT GATGGAATTG ATGCAGATGG
420
TACAGTGACA AAGGGAGGAT ATTCTAGTCA CATTGTCGTC CATGAAAGGT ATTGCGTCAG
480
GATTCCAGAA AACTACCCGA TGGATCTAGC AGCGCATTGC TCTGTGCTGG ATCAC
535

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 671 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCGCCTGCAG GTCGACACTA GTGGATCCAA AGAATTCGGC ACGAGGTTGC AGGTCGGGGA
60
TGATTTGAAT CACAGAAACC TCAGCGATTT TGCCAAGAAA TATGGCAAAA TCTTTCTGCT
120
CAAGATGGGC CAGAGGAATC TTGTGGTAGT TTCATCTCCC GATCTCGCCA AGGAGGTCCT
180
GCACACCCAG GGCGTCGAGT TTGGGTCTCG AACCCGGAAC GTGGTGTTCG ATATCTTCAC
240
GGGCAAGGGG CAGGACATGG TGTTCACCGT CTATGGAGAT CACTGGAGAA AGATGCGCAG
300
GATCATGACT GTGCCTTTCT TTACGAATAA AGTTGTCCAG CACTACAGAT TCGCGTGGGA
360
AGACGAGATC AGCCGCGTGG TCGCGGATGT GAAATCCCGC GCCGAGTCTT CCACCTCGGG
420
CATTGTCATC CGTAGCGCCT CCAGCTCATG ATGTATAATA TTATGTATAG GATGATGTTC
480
GACAGGAGAT TCGAATCCGA GGACGACCCG CTTTTCCTCA AGCTCAAGGC CCTCAACGGA
540
GAGCGAAGTC GATTGGCCCA GAGCTTTGAG TACAATTATG GGGATTTCAT TCCCAGTCTT
600
AGGCCCTTCC TCAGAGGTTA TCACAGAATC TGCAATGAGA TTAAAGAGAA ACGGCTCTCT
660
CTTTTCAAGG A
671

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 940 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTTCAGGACA AGGGAGAGAT CAATGAGGAT AATGTTTTGT ACATCGTTGA GAACATCAAC
60
GTTGCAGCAA TTGAGACAAC GCTGTGGTCG ATGGAATGGG GAATAGCGGA GCTGGTGAAC
120
CACCAGGACA TTCAGAGCAA GGTGCGCGCA GAGCTGGACG CTGTTCTTGG ACCAGGCGTG
180
CAGATAACGG AACCAGACAC GACAAGGTTG CCCTACCTTC AGGCGGTTGT GAAGGAAACC
240

CTTCGTCTCC GCATGGCGAT CCCGTTGCTC GTCCCCCACA TGAATCTCCA CGACGCCAAG
300
CTCGGGGGCT ACGATATTCC GGCAGAGAGC AAGATCCTGG TGAACGCCTG GTGGTTGGCC
360
AACAACCCCG CCAACTGGAA GAACCCCGAG GAGTTCCGCC CCGAGCGGTT CTTCGAGGAG
420
GAGAAGCACA CCGAAGCCAA TGGCAACGAC TTCAAATTCC TGNCCTTCGG TGTGGGGAGG
480
AGGAGCTGCC CGGGAATCAT TCTGGCGCTG CTCTCCTCGC ACTCTCCATC GGAAGACTTG
540
TTCAGAACTT CCACCTTCTG CCGCCGCCCG GGCAGAGCAA AGTGGATGTC ACTGAGAAGG
600
GCGGGCAATT CAGCCTTCAC ATTCTCAACC ATTCTCTCAT CGTCGCCAAG CCCATAGCTT
660
CTGCTTAATC CCAACTTGTC AGTGACTGGT ATATAAATGC GCGCACCTGA ACAAAAAACA
720
CTCCATCTAT CATGACTGTG TGTGCGTGTC CACTGTCGAG TCTACTAAGA GCTCATAGCA
780
CTTCAAAAGT TTGCTAGGAT TTCAATAACA GACACCGTCA ATTATGTCAT GTTTCAATAA
840
AAGTTTGCAT AAATTAAATG ATATTTCAAT ATACTATTTT GACTCTCCAC CAATTGGGGA
900
ATTTTACTGC TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
940

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 949 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

NNGCTCNACC GACGGTGGAC GGTCCGCTAC TCAGTAACTG AGTGGGATCC CCCGGGCTGA
60
CAGGCAATTC GATTTAGCTC ACTCATTAGG CACCCCAGGC TTTACACTTT ATGCTTCCGG
120
CTCGTATGTT GTGTGGAATT GTGAGCGGAT AACAATTTCA CACAGGAAAC AGCTATGACC
180
ATGATTACGC CAAGCGCGCA ATTAACCCTC ACTAAAGGGA ACAAAAGCTG GAGCTCCACC
240
GCGGTGGCGG CCGCTCTAGA ACTAGTGGAT CCAAAGAATT CGGCACGAGA CCCAGTGACC
300
TTCAGGCCTG AGAGATTTCT TGAGGAAGAT GTTGATATTA AGGGCCATGA TTACAGGCTA
360
CTGCCATTGG TGCAGGGCGC AGGATCTGCC CTGGTGCACA ATTGGGTATT AATTTAGTTC
420
AGTCTATGTT GGGACACCTG CTTCATCATT TCGTATGGGC ACCTCCTGAG GGAATGAAGG
480
CAGAAGACAT AGATCTCACA GAGAATCCAG GGCTTGTTAC TTTCATGGCC AAGCCTGTGC
540
AGGCCATTGC TATTCCTCGA TTGCCTGATC ATCTCTACAA GCGACAGCCA CTCAATTGAT
600
CAATTGATCT GATAGTAAGT TTGAATTTTG TTTTGATACA AAACGAAATA ACGTGCAGTT
660
TCTCCTTTTC CATAGTCAAC ATGCAGCTTT CTTTCTCTGA AGCGCATGCA GCTTTCTTTC
720
TCTGAAGCCC AACTTCTAGC AAGCAATAAC TGTATATTTT AGAACAAATA CCTATTCCTC
780
AAATTGAGWA TTTCTCTGTA GGGGNNGNTA ATTGTGCAAT TTGCAAGNAA TAGTAAAGTT
840
TANTTTAGGG NATTTTAATA GTCCTANGTA ANANGNGGNA ATGNTAGNGG GCATTNAGAA
900

ANCCCTAATA GNTGTTGGNG GNNGNTAGGN TTTTTNACCA AAAAAAAAA
949

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 959 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAATTCGGCA CGAGAAAGCC CTAGAATTTT TTCAGCATGC TATCACAGCC CCAGCGACAA
60
CTTTAACTGC AATAACTGTG GAAGCGTACA AAAAGTTTGT CCTAGTTTCT CTCATTCAGA
120
CTGGTCAGGT TCCAGCATTT CCAAAATACA CACCTGCTGT TGTCCAAAGA AATTTGAAAT
180
CTTGCACTCA GCCCTACATT GATTTAGCAA ACAACTACAG TAGTGGGAAA ATTTCTGTAT
240
TGGAAGCTTG TGTCAACACG AACACAGAGA AGTTCAAGAA TGATAGTAAT TTGGGGTTAG
300
TCAAGCAAGT TTTGTCATCT CTTTATAAAC GGAATATTCA GAGATTGACA CAGACATATC
360
TGACCCTCTC TCTTCAAGAC ATAGCAAGTA CGGTACAGTT GGAGACTGCT AAGCAGGCTG
420
AACTCCATGT TCTGCAGATG ATTCAAGATG GTGAGATTTT TGCAACCATA AATCAGAAAG
480
ATGGGATGGT GAGCTTCAAT GAGGATCCTG AACAGTACAA AACATGTCAG ATGACTGAAT
540
ATATAGATAC TGCAATTCGG AGAATCATGG CACTATCAAA GAAGCTCACC ACAGTAGATG
600
AGCAGATTTC GTGTGATCAT TCCTACCTGA GTAAGGTGGG GAGAGAGCGT TCAAGATTTG
660
ACATAGATGA TTTTGATACT GTTCCCCAGA AGTTCANAAA TATGTAACAA ATGATGTAAA
720
TCATCTTCAA GACTCGCTTA TATTCATTAC TTTCTATGTG AATTGATAGT CTGTTAACAA
780
TAGTACTGTG GCTGAGTCCA GAAAGGATCT CTCGGTATTA TCACTTGACA TGCCATCAAA
840
AAAATCTCAA ATTTCTCGAT GTCTAGTCTT GATTTTGATT ATGAATGCGA CTTTTAGTTG
900
TGACATTTGA GCACCTCGAG TGAACTACAA AGTTGCATGT TAAAAAAAAA AAAAAAAAA
959

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1026 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAATTCGGCA CGAGCTTTGA GGCAACCTAC ATTCATTGAA TCCCAGGATT TCTTCTTGTC
60
CAAACAGGTT TAAGGAAATG GCAGGCACAA GTGTTGCTGC AGCAGAGGTG AAGGCTCAGA
120
CAACCCAAGC AGAGGAGCCG GTTAAGGTTG TCCGCCATCA AGAAGTGGGA CACAAAAGTC
180
TTTTGCAGAG CGATGCCCTC TATCAGTATA TATTGGAAAC GAGCGTGTAC CCTCGTGAGC
240

CCGAGCCAAT GAAGGAGCTC CGCGAAGTGA CTGCCAAGCA TCCCTGGAAC CTCATGACTA
300
CTTCTGCCGA TGAGGGTCAA TTTCTGGGCC TCCTGCTGAA GCTCATTAAC GCCAAGAACA
360
CCATGGAGAT TGGGGTGTAC ACTGGTTACT CGCTTCTCAG CACAGCCCTT GCATTGCCCG
420
ATGATGGAAA GATTCTAGCC ATGGACATCA ACAGAGAGAA CTATGATATC GGATTGCCTA
480
TTATTGAGAA AGCAGGAGTT GCCCACAAGA TTGACTTCAG AGAGGGCCCT GCTCTGCCAG
540
TTCTGGACGA ACTGCTTAAG AATGAGGACA TGCATGGATC GTTCGATTTT GTGTTCGTGG
600
ATGCGGACAA AGACAACTAT CTAAACTACC ACAAGCGTCT GATCGATCTG GTGAAGGTTG
660
GAGGTCTGAT TGCATATGAC AACACCCTGT GGAACGGATC TGTGGTGGCT CCACCCGATG
720
CTCCCCTGAG GAAATATGTG AGATATTACA GAGATTTCGT GATGGAGCTA AACAAGGCCC
780
TTGCTGTCGA TCCCCGCATT GAGATCAGCC AAATCCCAGT CGGTGACGGC GTCACCCTTT
840
GCAGGCGTGT CTATTGAAAA CAATCCTTGT TTCTGCTCGT CTATTGCAAG CATAAAGGCT
900
CTCTGATTAT AAGGAGAACG CTATAATATA TGGGGTTGAA GCCATTTGTT TTGTTTAGTG
960
TATTGATAAT AAAGTAGTAC AGCATATGCA AAGTTTGTAT CAAAAAAAAA AAAAAAAAAA
1020
AAAAAA
1026

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1454 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAATTCGGCA CGAGGCCAAC TGCAAGCAAT ACAGTACAAG AGCCAGACGA TCGAATCCTG
60
TGAAGTGGTT CTGAAGTGAT GGGAAGCTTG GAATCTGAAA AAACTGTTAC AGGATATGCA
120
GCTCGGGACT CCAGTGGCCA CTTGTCCCCT TACACTTACA ATCTCAGAAA GAAAGGACCT
180
GAGGATGTAA TTGTAAAGGT CATTTACTGC GGAATCTGCC ACTCTGATTT AGTTCAAATG
240
CGTAATGAAA TGGACATGTC TCATTACCCA ATGGTCCCTG GGCATGAAGT GGTGGGGATT
300
GTAACAGAGA TTGGCAGCGA GGTGAAGAAA TTCAAAGTGG GAGAGCATGT AGGGGTTGGT
360
TGCATTGTTG GGTCCTGTCG CAGTTGCGGT AATTGCAATC AGAGCATGGA ACAATACTGC
420
AGCAAGAGGA TTTGGACCTA CAATGATGTG AACCATGACG GCACACCTAC TCAGGGCGGA
480
TTTGCAAGCA GTATGGTGGT TGATCAGATG TWTGTGGTTC GAATCCCGGA GAATCTTCCT
540
CTGGAACAAG CGGCCCCTCT GTTATGTGCA GGGGTTACAG TTTTCAGCCC AATGAAGCAT
600
TTCGCCATGA CAGAGCCCGG GAAGAAATGT GGGATTTTGG GTTTAGGAGG CGTGGGGCAC
660
ATGGGTGTCA AGATTGCCAA AGCCTTTGGA CTCCACGTGA CGGTTATCAG TTCGTCTGAT
720
AAAAAGAAAG AAGAAGCCAT GGAAGTCCTC GGCGCCGATG CTTATCTTGT TAGCAAGGAT
780

ACTGAAAAGA TGATGGAAGC AGCAGAGAGC CTAGATTACA TAATGGACAC CATTCCAGTT
840
GCTCATCCTC TGGAACCATA TCTTGCCCTT CTGAAGACAA ATGGAAAGCT AGTGATGCTG
900
GGCGTTGTTC CAGAGTCGTT GCACTTCGTG ACTCCTCTCT TAATACTTGG GAGAAGGAGC
960
ATAGCTGGAA GTTTCATTGG CAGCATGGAG GAAACACAGG AAACTCTAGA TTTCTGTGCA
1020
GAGAAGAAGG TATCATCGAT GATTGAGGTT GTGGGCCTGG ACTACATCAA CACGGCCATG
1080
GAAAGGTTGG AGAAGAACGA TGTCCGTTAC AGATTTGTGG TGGATGTTGC TAGAAGCAAG
1140
TTGGATAATT AGTCTGCAAT CAATCAATCA GATCAATGCC TGCATGCAAG ATGAATAGAT
1200
CTGGACTAGT AGCTTAACAT GAAAGGGAAA TTAAATTTTT ATTTAGGAAC TCGATACTGG
1260
TTTTTGTTAC TTTAGTTTAG CTTTTGTGAG GTTGAAACAA TTCAGATGTT TTTTTAACTT
1320
GTATATGTAA AGATCAATTT CTCGTGACAG TAAATAATAA TCCAATGTCT TCTGCCAAAT
1380
TAATATATGT ATTCGTATTT TTATATGAAA AAAAAAAAAA AAAA
1454AAAAAA AAAAAAAAAA 1440 AAAAAAAAAA AAAA
1454

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 740 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAATTCGGCA CGAGACCATT TCCAGCTAAT ATTGGCATAG CAATTGGTCA TTCTATCTTT
60
GTCAAAGGAG ATCAAACAAA TTTTGAAATT GGACCTAATG GTGTGGAGGC TAGTCAGCTA
120
TACCCAGATG TGAAATATAC CACTGTCGAT GAGTACCTCA GCAAATTTGT GTGAAGTATG
180
CGAGATTCTC TTCCACATGC TTCAGAGATA CATAACAGTT TCAATCAATG TTTGTCCTAG
240
GCATTTGCCA AATTGTGGGT TATAATCCTT CGTAGGTGTT TGGCAGAACA GAACCTCCTG
300
TTTAGTATAG TATGACGAGC TAGGCACTGC AGATCCTTCA CACTTTTCTC TTCCATAAGA
360
AACAAATACT CACCTGTGGT TTGTTTTCTT TCTTTCTGGA ACTTTGGTAT GGCAATAATG
420
TCTTTGGAAA CCGCTTAGTG TGGAATGCTA AGTACTAGTG TCCAGAGTTC TAAGGGAGTT
480
CCAAAATCAT GGCTGATGTG AACTGGTTGT TCCAGAGGGT GTTTACAACC AACAGTTGTT
540
CAGTGAATAA TTTTGTTAGA GTGTTTAGAT CCATCTTTAC AAGGCTATTG AGTAAGGTTG
600
GTGTTAGTGA ACGGAATGAT GTCAAATCTT GATGGGCTGA CTGACTCTCT TGTGATGTCA
660
AATCTTGATG GATTGTGTCT TTTTCAATGG TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
740 720 AAAAAAAAAA AAAAAAAAAA
740

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 624 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAATTCCTGC AGCCCGGGGG ATCCACTAGT TCTAGAGCGG CCGCCACCGC GGTGGAGCTC
60
GCGCGCCTGC AGGTCGACAC TAGTGGATCC AAAGAATTCG GCACGAGGCC CGACGGCCAC
120
TTGTTGGACG CCATGGAAGC TCTCCGGAAA GCCGGGATTC TGGAACCGTT TAAACTGCAG
180
CCCAAGGAAG GACTGGCTCT CGTCAACGGC ACAGCGGTGG GATCCGCCGT GGCCGCGTCC
240
GTCTGTGTTG ACGCCAACGT GCTGGGCGTG CTGGCTGAGA TTCTGTCTGC GCTCTTCTGC
300
GAGGTGATGC AAGGGAAACC GGAGTTCGTA GATCCGTTAA CCCACCAGTT GAAGCACCAC
360
CCAGGGCAGA TCGAAGCCGC GGCCGTCATG GAGTTCCTCC TCGACGGTAG CGACTACGTG
420
AAAGAAGCAG CGCGGCTTCA CGAGAAAGAC CCGTTGAGCA AACCGAAACA AGACCGCTAC
480
GCTCTGCGAA CATCGCCACA GTGGTTGGGG CCTCCGATCG AAGTCATCCG CGCTGCYACT
540
CACTCCATCG AGCGGGAGAT CAATTCCGTC AACGACAATC CGTTAATCGA TGTCTCCAGG
600
GACATGGCTG TCCACGGCGG CAAC
624

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 278 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAATTCCTGC AGCCCGGGGG ATCCACTAGT TCTAGAGCGG CCGCCACCGC GGTGGAGCTC
60
CAGTACCTGG CCAACCCCGT CACGACTCAC GTCCAGAGCG CCGAACAACA CAACCAGGAT
120
GTCAATTCCC TCGGCTTGAT CTCCGCCAGA AAGACTGCCG AGGCCGTTGA GATTTTAAAG
180
CTGATGTTCG CTACATATCT GGTGGCCTTA TGCCAGGCGA TCGATCTCCG GCACCTGGAA
240
GAAAACATGC GATCCGTTGT GAAGCACGTA GTCTTGCA
278

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 765 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAGCTCCTGC AAGTCATCGA TCATCAGCCC GTTTTCTCGT ACATCGACGA TCCCACAAAT
60
CCATCATACG CGCTTATGCT CCAACTCAGA GAAGTGCTCG TAGATGAGGC TCTCAAATCA
120

TCTTGCCCAG ACGGGAATGA CGAATCCGAT CACAATTTGC AGCCCGCTGA GAGCGCTGGA
180
GCTGCTGGAA TATTACCCAA TTGGGTGTTT AGCAGGATCC CCATATTTCA AGAGGAGTTG
240
AAGGCCCGTT TAGAGGAAGA GGTTCCGAAG GCGAGGGAAC GATTCGATAA TGGGGACTTC
300
CCAATTGCAA ACAGAATAAA CAAGTGCAGG ACATATCCCA TTTACAGATT CGTGAGATCA
360
GAGTTGGGAA CCGATTTGCT AACAGGGCCC AAGTGGAGAA GCCCCGGCGA AGATATAGAA
420
AAGGTATTTG AGGGCATTTG CCAAGGGAAA ATTGGAAACG TGATCCTCAA ATGTCTGGAC
480
GCTTGGGGTG GGTGCGCTGG ACCATTCACT CCACGTGCAT ATCCTGCGTC TCCTGCAGCG
540
TTCAATGCCT CATATTGGGC ATGGTTTGAT AGCACCAAAT CACCCTCTGC AACGAGCGGC
600
AGAGGTTTCT GGAGCGCCCA ACAACAACAA GTTCTTTGAT TTAACTGACT CTTAAGCATT
660
CCTAAACAGC TTGTTCTTCG CAATAACGAA TCTTTCATCT TCGTTACTTT GTAAAAGATG
720
GGGTTCCAAC AAAATAGAAG AAATATTTTC GATCCAAAAA AAAAA
765

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 453 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TGATTATGCG GATCCTTGGG CAGGGATACG GCATGACAGA AGCAGGCCCG GTGCTGGCAA
60
TGAACCTAGC CTTCGCAAAG AATCCTTTCC CCGCCAAATC TGGCTCCTGC GGAACAGTCG
120
TCCGGAACGC TCAAATAAAG ATCCTCGATT ACAGGAACTG GCGAGTCTCT CCCGCACAAT
180
CAAGCCGGCG AAATCTGCAT CCGCGGACCC GAAATAATGA AAGGATATAT TAACGACCCG
240
GAATCCACGG CCGCTACAAT CGATGAAGAA GGCTGGCTCC ACACAGGCGA CGTCGGGTAC
300
ATTGACGATG ACGAAGAAAT CTTCATAGTC GACAGAGTAA AGGAGATTAT CAATATAAAG
360
GCTTCCAGGT GGATCCTGCT AATCGAATTC CTGCAGCCCG GGGGTCCACT AGTTCTAGAG
420
CGGCCGCCAC CGCGGTGGAG CTCCAGCTTT TGT
453

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 278 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCTTCGAATT CTCTTTCACG ACTGCTTCGT TAATGGCTGC GATGGCTCGA TATTGTTAGA
60
TGATAACTCA ACGTTCACCG GAGAAAAGAC TGCAGGCCCA AATGTTAATT CTGCGAGAGG
120

ATTCGACGTA ATAGACACCA TCAAAACTCA AGTTGAGGCA GCCTGCAGTG GTGTCGTGTC
180
AGTTGCCGAC ATTCTCGCCA TTGCTGCACG CGATTCAGTC GTCCAACTGG GGGGCCCAAC
240
ATGGACGGTA CTTCTGGGAG AAAAGACGGA TCCGATCA
278

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTTCGAATTC WYTTYCAYGA YTG
23

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GATCGGATCC RTCYYKYCTY CC
22

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AATTCGGCAC GAGACGACCT CTTGTATCGG ACCCGGATCC GCTATCGTTA ACGTACACAC
60
GTTCTAGTGC TGAATGGAGA TGGAGAGCAC CACCGGCACC GGCAACGGCC TTCACAGCCT
120
CTGCGCCGCC GGGAGCCACC ATGCCGACCC ACTGAACTGG GGGGCGGCGG CAGCAGCCCT
180
CACAGGGAGC CACCTCGACG AGGTGAAGCG GATGGTCGAG GAGTACCGGA GGCCGGCGGT
240
GCGCCTCGGC GGGGAGTCCC TCACGATAGC CCAGGTGGCG GCGGTGGCGA GTCAGGAGGG
300
GGTAGGGGTC GAGCTCTCGG AGGCGGCCCG TCCCAGGGTC AAGGCCAGCA GCGACTGGGT
360
CATGGAGAGC ATGAACAAGG GAACTGACAG CTACGGGGTC ACCACCGGGT TCGGCGGCAA
420
CTTCTCAAAC CGGAGGCCGA AGCAAGGCGG TCCTTTTCAG AAGGAACTTA TA
472

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 622 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCAAAGCTCC TAGTGCCTCA TGAGTCTGCT GAGGATTGCA CAATTGGCGG GTTCGACGTG
60
CCCCGAGGCA CCATGATCCT GGTTAATGCG TGGGCAATTC AAAGAGACCC AAAAGTGTGG
120
GACGATCCCA CAAATTTTAA ACCGGAGAGG TACGAGGGAT TGGAAGGTGA TCATGCCTAC
180
CGACTATTGC CGTTTGGGAT GGGGAGGAGA AGTTGTCCTG GTGCTGGCCT TGCCAATAGA
240
GTGGTGAGCT TGGTCCTGGC GGCGCTTATT CAGTGCTTCG AATGGGAACG AGTTGGCGAA
300
GAATTGGTGG ACTTGTCCGA GGGGACGGGA CTCACAATGC CAAAGAGAGA GCCATTGGAG
360
GCCTTGTGCA AAGCGCGTGA ATGCATGATA GCTAATGTTC TTGCGCACCT TTAAGAAGGT
420
CGTTGTCTAA TGAATTTACA TTGGTGATGT ATCTCCAATG TTTTTGAATA ATCAAATAGA
480
CTGAAAATAG GCCAGTGCAG CTTTAGGAAT GATCGTGAGC ATCAATAGCA TCCTGAGGAG
540
GCCAATGCAG CTTTAGGCCT TTCTCTTAGG AGAAAAATGA TGGTTTATAT AGGTACTGGC
600
AACATTGTTC AAAAAAAAAA AA
622

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 414 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CACGCTCGAC GAATTCGGTA CCCCGGGTTC GAAATCGATA AGCTTGGATC CAAAGCAACA
60
CATTGAACTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCCCCCACCC CCCCTTCCCA
120
ACCCCACCCA CATACAGACA AGTAGATACG CGCACACAGA AGAAGAAAAG ATGGGGGTTT
180
CAATGCAGTC AATCGCACTA GCGACGGTTC TGGCCGTCCT AACGACATGG GCGTGGAGGG
240
CGGTGAACTG GGTGTGGCTG AGGCCGAAGA GGCTCGAGAG GCTTCTGAGA CAGCAAGGTC
300
TCTCCGGCAA GTCCTACACC TTCCTGGTCG GCGACCTCAA GGAGAACCTG CGGATGCTCA
360
AGGAAGCCAA GTCCAAGCCC ATCGCCGTCT CCGATGACAT CAAGCCTCGT CTCT
414

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 469 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAATTCGGCA CGAGTGTCTC TCTCTCTCTC TCTCTCTGTA AACCACCATG CTCTTCCTCA
60
CTCATCTCCT AGCAGTTCTA GGGGTTGTGT TGCTCCTGCT AATTCTATGG AGGGCAAGAT
120
CTTCTCCGAA CAAACCCAAA GGTACTGCCT TACCCCCGGA GCTGCCGGGC GCATGGCCGA
180
TCATAGGCCA CATCCACTTG CTGGGCGGCG AGACCCCGCT GGCCAGGACC CTGGCCGCCA
240
TGGCGGACAA GCAGGGCCCG ATGTTTCGGA TCCGTCTCGG AGTCCACCCG GCGACCATCA
300
TAAGCAGCCG TGAGGCGGTC CGGGAGTGCT TCACCACCCA CGACAAGGAC CTCGCTTCTC
360
GCCCCAAATC CAAGGCGGGA ATCCACTTGG GCTACGGGTA TGCCGGTTTT GGCTTCGTAG
420
AATACGGGGA CTTTTGGCGC GAGATGAGGA AGATCACCAT GCTCGAGCT
469

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 341 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGGCTCGTG GCTCGGCTCC GGCGCAACGC CCTTCCCACC GGGCCCGAGG GGCCTCCCGG
60
TCATCGGGAA CATGCTCATG ATGGGCGAGC TCACCCACCG CGGCCTCGCG AGTCTGGCGA
120
AGAAGTATGG CGGGATCTTC CACCTCCGCA TGGGCTTCCT GCACATGGTT GCCGTGTCGT
180
CCCCCGACGT GGCCCGCCAG GTCCTCCAGG TCCACGACGG GATCTTCTCG AACCGGCCTG
240
CCACCATCGC GATCAGCTAC CTCACGTATG ACCGGGCCGA CATGGCCTTC GCGCACTACG
300
GCCCGTTCTG GCGGCAGATG CGGAAGCTGT GCGTGATGAA A
341

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 387 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GAATTCGGCA CGAGCGGGCT CGTGGCTCGG CTCCGGCGCA ACGCCCTTCC CACCGGGCCC
60
GAGGGGCCTC CCGGTCATCG GGAACATGCT CATGATGGGC GAGCTCACCC ACCGCGGCCT
120
CGCGAGTCTG GCGAAGAAGT ATGGCGGGAT CTTCCACCTC CGCATGGGCT TCCTGCACAT
180
GGTTGCCGTG TCGTCCCCCG ACGTGGCCCG CCAGGTCCTC CAGGTCCACG ACGGGATCTT
240
CTCGAACCGG CCTGCCACCA TCGCGATCAG CTACCTCACG TATGACCGGG CCGACATGGC
300
CTTCGCGCAC TACGGCCCGT TCTGGCGGCA GATGCGGAAG CTGTGCGTGA TGAAAGCTCT
360

TCAGCGGAAG CGGGCTGAGT CGTGGGA
387

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 443 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CACGAGCTCG TGAGCCTTCC CGGAGACAAG GCCATCTTAC TTCGCAACAA ATTGCGTCCG
60
CACTCCTTTC TCAAGAAACC TAGTCATCCA AGAAGCAGAG CATTGCAACT GCAAACAGCC
120
AAAGCCCAAA CTCGTACAGA AGGAGAGAGA GAGAGAGAAT AGAAGCATGA GTGCATGCAC
180
GAACCAAGCA ATCACGACGG CCAGTGAAGA TGAAGAGTTC TTGTTCGCCA TGGAAATGAA
240
TGCTCTGATA GCACTCCCCT TGGTCTTGAA GGCCACCATC GAACTGGGGA TCCTCGAAAT
300
ACTGGCCGAG TGCGGGCCTA TGGCTCCACT TTCGCCTGCT CAGATTGCCT CCCGTCTCTC
360
CGCAAAGAAC CCGGAAGCCC CCGTAACCCT TGACCGGATC CTCCGGTTTC TCGCCAGCTA
420
CTCCATCCTC TCTTGCACTC TCG
443

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 607 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GAATTCGGCA CGAGCCAACC CTGGACCAGG TACTTTTGGC AGGCGGTCCA TTGCCCTTCA
60
AACCGGTCCA AACCGGACCA TCACTGTCCT TATATACGTT GCATCATGCC TGCTCATAGA
120
ACTTAGGTCA ACTGCAACAT TTCTTGATCA CAACATATTA CAATATTCCT AAGCAGAGAG
180
AGAGAGAGAG AGAGAGAGAG AGAGAGAGAG AGAGTTTGAA TCAATGGCCA CCGCCGGAGA
240
GGAGAGCCAG ACCCAAGCCG GGAGGCACCA GGAGGTTGGC CACAAGTCTC TCCTTCAGAG
300
TGATGCTCTT TACCAATATA TTTTGGAGAC CAGCGTGTAC CCAAGAGAGC CTGAGCCCAT
360
GAAGGAGCTC AGGGAAATAA CAGCAAAACA TCCATGGAAC ATAATGACAA CATCAGCAGA
420
CGAAGGGCAG TTCTTGAACA TGCTTCTCAA GCTCATCAAA GCCAAGAACA CCATGGAGAT
480
TGGTGTCTTC ACTGGCTACT CTCTCCTCGC CACCGCTCTT GCTCTTCCTG ATGACGGAAA
540
GATTTTGGCT ATGGACATTA ACAGAGAGAG CTATGAACTT GGCCTGCCGG CATCCAAAAA
600
GCCGGTG
607

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 421 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAATTCGGCA CGAGCCGTTT TATTTCCTCT GATTTCCTTT GCTCGAGTCT CGCGGAAGAG
60
AGAGAAGAGA GGAGAGGAGA GAATGGGTTC GACCGGATCC GAGACCCAGA TGACCCCGAC
120
CCAAGTCTCG GACGAGGAGG CGAACCTCTT CGCCATGCAG CTGGCGAGCG CCTCCGTGCT
180
CCCCATGGTC CTCAAGGCCG CCATCGAGCT CGACCTCCTC GAGATCATGG CCAAGGCCGG
240
GCCGGGCGCG TTCCTCTCCC CGGGGGAAGT CGCGGCCCAG CTCCCGACCC AGAACCCCGA
300
GGCACCCGTA ATGCTCGACC GGATCTTCCG GCTGCTGGCC AGCTACTCCG TGCTCACGTG
360
CACCCTCCGC GACCTCCCCG ATGGCAAGGT CGAGCGGCTC TACGGCTTAG CGCCGGTGTC
420
C
421

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 760 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGAAGAAGCC GAGCAAACGA ATTGCAGACG CCATTGAAAA AAGACACGAA AGAGATCAAG
60
AAGGAGCTTA AGAAGCATCA TCAATGGCAG CCAACGCAGA GCCTCAGCAG ACCCAACCAG
120
CGAAGCATTC GGAAGTCGGC CACAAGAGCC TCTTGCAGAG CGATGCTCTC TACCAGTATA
180
TATTGGAGAC CAGCGTCTAC CCAAGAGAGC CAGAGCCCAT GAAGGAGCTC AGGGAAATAA
240
CAGCCAAACA TCCATGGAAC CTGATGACCA CATCGGCGGA TGAAGGGCAG TTCCTGAACA
300
TGCTCCTCAA GCTCATCAAC GCCAAGAACA CCATGGAGAT CGGCGTCTAC ACCGGCTACT
360
CTCTCCTCGC AACCGCCCTT GCTCTTCCCG ATGACGGAAA GATCTTGGCC ATGGCCATCA
420
ATAGGGAGAA CTTCGAGATC GGGCTGCCCG TCATCCAGAA GGCCGGCCTT GCCCACAAGA
480
TCGATTTCAG AGAAGGCCCT GCCCTGCCGC TCCTTGATCA GCTCGTGCAA GATGAGAAGA
540
ACCATGGAAC GTACGACTTC TTCTCAATCC TTAATCGTTC ATTTGAATAC AAATACATGC
600
TCAATGGTTC AAAGACAACA TAAGACAGAA GATGGAAAAA ATAGAAAGGA AGGAAAGTAT
660
TAAGGGTAGT TTCTCATTTC ATCAATGCTT GATTTTGAGA TCTCCTTTCT GGTGCGATCA
720
GCTGACCCGG CGGCACAGGT GATGCCATCC CCGACGGGAA
760

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 508 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GAATTCGGTA CCCGGGTTCG AAATCGATAA GCTTGGATCC AAAGAATTCG GCACGAGATC
60
ACTAACCATC TGCCTTTCTT CATCTTCTTT CTTCTGCTTC TCCTCCGTTT CCTCGTTTCG
120
ATATCGTGAA AGGAGTCCGT CGACGACAAT GGCCGAGAAG AGCAAGGTCC TGATCATCGG
180
AGGGACGGGC TACGTCGGCA AGTTCATCGT GGAAGCGAGT GCAAAAGCAG GGCATCCCAC
240
GTTCGCGCTG GTTAGGCAGA GCACGGTCTC CGACCCCGTC AAGGGCCAGC TCGTCGAGAG
300
CTTCAAGAAC TTGGGCGTCA CTCTGCTCAT CGGTGATCTG TACGATCATG AGAGCTTGGT
360
GAAGGCAATC AAGCAAGCCG ACGTGGTGAT ATCGACAGTG GGGCACATGC AAATGGCGGA
420
TCAGACCAAA GAATCGTCGA CGCCATTAAA GGAAGCTGGC AACGTTAAGG TTTGTTGGTT
480
GGTTCATTTG ATCTGGTTTG GGGGGGTC
508

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 495 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAATTCGGCA CGAGGTTAAT GGCAGTGCAG CCTCAACACC ACCCACCTTC CTCCATCTCT
60
CTCCTCCCTT CTTCTTTCTC TGACTTCAAT GGCAGCCGAC TCCATGCTTG CGTTCAGTAT
120
AAGAGGAAGG TGGGGCAGCC TAAAGGGGCA CTGCGGGTCA CTGCATCAAG CAATAAGAAG
180
ATCCTCATCA TGGGAGGCAC CCGTTTCATC GGTGTGTTTT TGTCGAGACT ACTTGTCAAA
240
GAAGGTCATC AGGTCACTTT GTTTACCAGA GGAAAAGCAC CCATCACTCA ACAATTGCCT
300
GGTGAGTCGG ACAAGGACTT CGCTGATTTT TCATCCAAGA TCCTGCATTT GAAAGGAGAC
360
AGAAAGGATT TTGATTTTGT TAAATCTAGT CTTGCTGCAG AAGGCTTTGA CGTTGTTTAT
420
GACATTAACG GCGAGAGGCG GATGAAGTCG CACCAATTTT GGATGCCTGC CAAACCTTGA
480
ACCAGTCAAC TACTG
495

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GAATTCGGCA CGAGCATAAG CTCTCCCGTA ATCCTCACAT CACATGGCGA AGAGCAAGGT
60
CCTCGTCGTT GGCGGCACTG GCTACCTCGG GCGGAGGTTC GTGAGGGCGA GCCTGGACCA
120
GGGCCACCCC ACGTACGTCC TCCAGCGTCC GGAGACCGGC CTCGACATTG AGAAGCTCCA
180
GACGCTACTG CGCTTCAAGA GGCGTGGCGC CCAACTCGTC GAGGCCTCGT TCTCAGACCT
240
GAGGAGCCTC GTCGACGCTG TGAGGCGGGT CGATGTCGTC GTCTGTGCCA TGTCGGGGGT
300
CCACTTCCGG AGCCACAACA TCCTGATGCA GCTCAAGCTC GTGGAGGCTA TCAAAGAAGC
360
TGGAAATGTC AAGCGGTTTT TGCCGTCAGA GTTCGGAATG GACCCGGCCC TCATGGGTCA
420
TGCAATTGAG CCGGGAAGGG TCACGTTCGA TGAGAAATGG AGGTGAGAAA AG
472

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 396 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GAATTCGGCA CGAGGAGGCA CCTCCTCGAA ACGAAGAAGA AGAAGGACGA AGGACGAAGG
60
AGACGAAGGC GAGAATGAGC GCGGCGGGCG GTGCCGGGAA GGTCGTGTGC GTGACCGGGG
120
CGTCCGGTTA CATCGCCTCG TGGCTCGTCA AGCTCCTCCT CCAGCGCGGC TACACCGTCA
180
AGGCCACCGT CCGCGATCCG AATGATCCAA AAAAGACTGA ACATTTGCTT GGACTTGATG
240
GAGCGAAAGA TAGACTTCAA CTGTTCAAAG CAAACCTGCT GGAAGAGGGT TCATTTGATC
300
CTATTGTTGA GGGTTGTGCA GGCGTTTTTC AAACTGCCTC TCCCTTTTAT CATGATGTCA
360
AGGATCCGCA GGCAGAATTA CTTGATCCGG CTGTAA
396

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 592 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GAATTCGGCA CGAGGTTGAA CCTCCCGTCC TCGGCTCTGC TCGGCTCGTC ACCCTCTTCG
60
CGCTCCCGCA TACTCCACCA CCGCGTACAG AAGATGAGCT CGGAGGGTGG GAAGGAGGAT
120
TGCCTCGGTT GGGCTGCCCG GGACCCTTCT GGGTTCCTCT CCCCCTACAA ATTCACCCGC
180
AGGGCCGTGG GAAGCGAAGA CGTCTCGATT AAGATCACGC ACTGTGGAGT GTGCTACGCA
240

GATGTGGCTT GGACTAGGAA TGTGCAGGGA CACTCCAAGT ATCCTCTGGT CCCAGGGCAC
300
GAGATAGTTG GAATTGTGAA ACAGGTTGGC TCCAGTGTCC AACGCTTCAA AGTTGGCGAT
360
CATGTGGGGG TGGGAACTTA TGTCAATTCA TGCAGAGAGT GCGAGTATTG CAATGACAGG
420
CTAGAAGTCC AATGTGAAAA GTCGGTTATG ACTTTTGATG GAATTGATGC AGATGGTACA
480
GTGACAAAGG GAGGATATTC TAGTCACATT GTCGTCCATG AAAGGTATTG CGTCAGGATT
540
CCAGAAAACT ACCCGATGGA TCTAGCAGCG CATTTGCTCT GTGCTGGATC AC
592

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GAATTCGGCA CGAGAACTCA TCTTGAAATG TCATTGGAGT CATCATCCTC TAGTGAGAAG
60
AAACAAATGG GTTCCGCCGG ATTCGAATCG GCCACAAAGC CGCACGCCGT TTGCATTCCC
120
TACCCTGCAC AAAGCCACAT TGGCGCCATG CTCAAGCTAG CAAAGCTCCT CCATCACAAG
180
GGCTTCCACA TCTCCTTCGT CAACACCGAG TTCAACCACC GGCGGCTCGC CAGGGCTCGA
240
GGCCCCGAGT TCACAAATGG AATGCTGAGC GACTTTCAGT TCCTGACAAT CCCCGATGGT
300
CTTCCTCCTT CGGACTTGGA TGCGATCCAA GACATCAAGA TGCTCTGCGA ATCGTCCAGG
360
AACTATATGG TCAGCCCCAT CAACGATCTT GTATCGAGCC TGGGCTCGAA CCCGAGCGTC
420
CCTCCGGTGA CTTGCATCAA TCTCGGATGG TTTCATGACA CTCGTGAC
468

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 405 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTTTACTCCG CCAAGAAGAT CCAATCGCAG TTTTCGCAAT TGGCCCATTA CACAAATGCG
60
GTCCATCTTC ATCGGGAAGT CTCTTGGCAG AAGACCGGAG TTGCATTTCC TGGCTGGACA
120
AGCAAGCCCC TAACTCAGTG GTCTATGTGA GTCTTGGGAG CATCGCCTCT GTGAACGAGT
180
CGGAATTTTC CGAAATAGCT TTAGGTTTAG CCGATAGCCA GCAGCCATTC TTGTGGGTGG
240
TTCGACCCGG GTCAGTGAGC GGCTCGGAAC TCTTAGAGAA TTTGCCCGGT TGCTTTCTGG
300
AGGCATTACA GGAGAGGGGG AAGATTGTGA AATGGGCGCC TCAACATGAA GTGCTGGCTC
360
ATCGGGCTGT CGGAGCGTTT TGGACTCACA ATGGATGGAA CTCCA
405

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 380 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GGCAAACACG CCCGTTTTCG TTTTACTAAG AGAAGATGGT GAGCGTTGTG GCTGGTAGAG
60
TCGAGAGCTT GTCGAGCAGT GGCATTCAGT CGATCCCGCA GGAGTATGTG AGGCCGAAGG
120
AGGAGCTCAC AAGCATTGGC GACATCTTCG AGGAGGAGAA GAAGCATGAG GGCCCTCAGG
180
TCCCGACCAT CGACCTCGAG GACATAGCGT CTAAAGACCC CGTGGTGAGG GAGAGGTGCC
240
ACGAGGAGCT CAGGAAGGCT GCCACCGACT GGGGCGTCAT GCACCTCGTC AACCATGGGA
300
TCCCCAACGA CCTGATTGAG CGTGTAAAGA AGGCTGGCGA GGTGTTCTTC AACCTCCCGA
360
TCGAGGAGAA GGACAAGCAT
380

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 305 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTGTACCCGA AGATCTCCGG GACCGTTCGA CGGCGACATC GCCGTCGGCC GGGAACCCGT
60
CGAGGCCGCC GCCGGAGGCC GGGGAGAAGC TGGAGTAGCC GCCGTAGCCG GAGAAGGCGC
120
CGTCGTGGTC GGCGGCGGCG GCGTGGTGGA CCTCATCGCC GTCCATGCTG AAGGCGTCGA
180
AGGAAGCGGA CATGGCTGGG GGATCGATCG ACCGATCCGA TCGGCCGGAG GATTTCGAGA
240
TCGGAGATGG AGAGATGGAA ATGAAAGAGA GAGAGAGAGA GAGATCCGGT GGACTGGTGG
300
TGTTT
305

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 693 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GAATTCGGCA CGAGCTAAGA GAGGAGAGGA GAGGAGCAAG ATGGCACTAG CAGGAGCTGC
60
ACTGTCAGGA ACCGTGGTGA GCTCCCCCTT TGTGAGGATG CAGCCTGTGA ACAGACTCAG
120

GGCATTCCCC AATGTGGGTC AGGCCCTGTT TGGTGTCAAC TCTGGCCGTG GCAGAGTGAC
180
TGCCATGGCC GCTTACAAGG TCACCCTGCT CACCCCTGAA GGCAAAGTCG AACTCGACGT
240
CCCCGACGAT GTTTACATCT TGGACTACGC CGAGGAGCAA GGCATCGACT TGCCCTACTC
300
CTGCCGTGCC GGCTCTTGCT CCTCCTGCGC GGGCAAGGTC GTGGCGGGGA GCGTCGACCA
360
GAGCGACGGC AGCTTCCTGG ATGATGATCA GATTGAGGAA GGTTGGGTCC TCACTTGTGT
420
CGCCTACCCT AAGTCTGAGG TCACCATTGA GACCCACAAG GAAGAGGAGC TCACTGCTTG
480
AAGCTCTCCT ATATTTGCTT TTGCATAAAT CAGTCTCACT CTACGCAACT TTCTCCACTC
540
TCTCCCCCCT TCACTACATG TTTGTTAGTT CCTTTAGTCT CTTCCTTTTT TACTGTACGA
600
GGGATGATTT GATGTTATTC TGAGTCTAAT GTAATGGCTT TTCTTTTTCC TATTTCTGTA
660
TGAGGAAATA AAACTCATGC TCTAAAAAAA AAA
693

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 418 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AGGACTTTAT TATAAGCATT GTAAAAAGAG TCAAACTAAT ACATCGCAAG AATTGGGTTA
60
TCCAATAATC TACAAAAAGA AAAAAGTTTG ATGCATTGAG ATGGTAACTG CTTAATTCAA
120
ATGCCTTAGT TTGAAAAATT AACCAACTAT TAAAATTAAT GATGATGAAT ATGGATTATG
180
TGTGAAAAAC TATATAGACT TAAAATTGAC TCAGAAGACA TTCTTTTCTT CTTATTTTAT
240
GATATGATGA ATTCGGTCTA AACAGGCAAA TGGTGTCAAA CGGGAAGTCG GCAAAACTCT
300
TCCTCGGCAG TGACTACCGG GCGGGCGATG ATGCGGATCC GGGGGCCGGG TCGCTGGAGA
360
ACATCCCGCA CGGACCGGTC CACGTTTGGT GCGGTGACAA CAGGCAGCCC AACCTGGA
418

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 777 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GAATTCGGCA CGAGCATACA ACTACACTGC GACGCCGCCG CAGAACGCGA GCGTGCCGAC
60
CATGAACGGC ACCAAGGTCT ACCGGTTGCC GTATAACGCT ACGGTCCAGC TCGTTTTACA
120
GGACACCGGG ATAATCGCGC CGGAGACCCA CCCCATCCAT CTGCACGGAT TCAACTTCTT
180
CGGTGTGGGC AAAGGAGTGG GGAATTATGA CCCAAAGAAG GATCCCAAGA AGTTCAATCT
240

GGTTGACCCA GTGGAGAGGA ACACCATTGG AATCCCATCT GGTGGATGGA TAGCCATCAG
300
ATTCACAGCA GACAATCCAG GAGTTTGGTT CCTGCACTGC CATCTGGAAG TGCACACAAC
360
TTGGGGACTG AAGATGGCAT TCTTGGTGGA CAATGGGAAG GGGCCTAAAG AGACCCTGCT
420
TCCACCTCCA AGTGATCTTC CAAAATGTTG ATCATTTGAT CATGAGGACG ACAAGCGATT
480
ACTAATGACA CCAAGTTAGT GGAATCTTCT CTTTGAAAAA GAAGAAGAAG AGCAAGAAGA
540
ATAAGAAAGA TGAGGAGAGA AGCCATAGAA GATTTGACCA AGAAGAGAGA GGGCAATAAA
600
CCAAAGAGAC CCTTGAGATC ACGACATCCC GCAATTGTTT CTAGAGTAAT AGAAGGATTT
660
ACTCCGACAC TGCTACAATA AATTAAGGAA GACAAGGAAT TTGGTTTTTT TCATTGGAGG
720
AGTGTAATTT GTTTTTTGGC AAGCTCATCA CATGAATCAC ATGGAAAAAA AAAAAAA
777

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 344 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

ATATGTTCAG AATTTCAAAT GTGGGAATGT CAACCTCCTT GAACTTCAGA ATTCAGGGCC
60
ATACGTTGAA GCTAGTCGAG GTTGAAGGAT CTCACACCGT CCAGAACATG TATGATTCAA
120
TCGATGTTCA CGTGGGCCAA TCCATGGCTG TCTTAGTGAC CTTAAATCAG CCTCCAAAGG
180
ACTACTACAT TGTCGCATCC ACCCGGTTCA CCAAGACGGT TCTCAATGCA ACTGCAGTGC
240
TACACTACAC CAACTCGCTT ACCCCAGTTT CCGGGCCACT ACCAGCTGGT CCAACTTACC
300
AAAAACATTG GTCCATGAAG CAAGCAAGAA CAATCAGGTG GAAC
344

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 341 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GCCGCAACTG CAATTCTCTT CGTAAAACAT GACGGCTGTC GGCAAAACCT CTTTCCTCTT
60
GGGAGCTCTC CTCCTCTTCT CTGTGGCGGT GACATTGGCA GATGCAAAAG TTTACTACCA
120
TGATTTTGTC GTTCAAGCGA CCAAGGTGAA GAGGCTGTGC ACGACCCACA ACACCATCAC
180
GGTGAACGGG CAATTCCCGG GTCCGACTTT GGAAGTTAAC GACGGCGACA CCCTCGTTGT
240
CAATGTCGTC AACAAAGCTC GCTACAACGT CACCATTCAC TGGCACGGCG TCCGGCAGGT
300
GAGATCTGGT TGGGCTGATG GGGCGGAATT TGTGACTCAA T
341

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 358 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GAATTCGGCA CGAGATATGT TCAGAATTTC AAATGTGGGA ATGTCAACCT CCTTGAACTT
60
CAGAATTCAG GGCCATACGT TGAAGCTAGT CGAGGTTGAA GGATCTCACA CCGTCCAGAA
120
CATGTATGAT TCAATCGATG TTCACGTGGG CCAATCCATG GCTGTCTTAG TGACCTTAAA
180
TCAGCCTCCA AAGGACTACT ACATTGTCGC ATCCACCCGG TTCACCAAGA CGGTTCTCAA
240
TGCAACTGCA GTGCTACACT ACACCAACTC GCTTACCCCA GTTTCCGGGC CACTACCAGC
300
TGGTCCAACT TACCAAAAAC ATTGGTCCAT GAAGCAAGCA AGAACAATCA GGTGGAAC
358

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 409 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

ATCAAGAGTT TGAGTCTAAA CCTTGTCTAA TCCTCTCTCG CATAGTCATT TGGAGACGAA
60
TGCTGATCGG CCGCAGCTGC ATTCTCTTCG TAAAACATGA CGGCTGTCGG CAAAACCTCT
120
TTCCTCTTGG GAGCTCTCCT CCTCTTCTCT GTGGCGGTGA CATTGGCAGA TGCAAAAGTT
180
TACTACCATG ATTTTGTCGT TCAAGCGACC AAGGTGAAGA GGCTGTGCAC GACCCACAAC
240
ACCATCACGG TGAACGGGCA ATTCCCGGGT CCGACTTTGG AAGTTAACGA CGGCGACACC
300
CTCGTTGTCA ATGTCGTCAA CAAAGCTCGC TACAACGTCA CCATTCACTG GCACGGCGTC
360
CGGCAGGTGA GATCTGGTTG GGCTGATGGG GCGGAATTTG TGACTCAAT
409

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTCTCTCTCT CTCTCTCTCT GTGTGTTCAT TCTCGTTGAG CTCGTGGTCG CCTCCCGCCA
60
TGGATCCGCA CAAGTACCGT CCATCCAGTG CTTTCAACAC TTCTTTCTGG ACTACGAACT
120

CTGGTGCTCC TGTCTGGAAC AATAACTCTT CGTTGACTGT TGGAAGCAGA GGTCCAATTC
180
TTCTTGAGGA TTATCACCTC GTGGAGAAAC TTGCCAACTT TGATAGGGAG AGGATTCCAG
240
AGCGTGTGGT GCATGCCAGA GGAGCCAGTG CAAAGGGATT CTTTGAGGTC ACTCATGACA
300
TTTCCCAGCT TACCTGTGCT GATTTCCTTC GGGCACCAGG AGTTCAAACA CCCGTGATTG
360
TCCGTTTCTC CACTGTCATC CACGAAAGGG GCAGCCCTGA AACCCTGAGG GACCCTCGAG
420
GTTTTGCTGT GAAGTTCTAC ACAAGAGAGG GTAACTTTGA TCTGGTGGGA AACAATTTCC
480
CTGTCTTCTT TGTCCGTAAT GGGATAAATT CCCCG
515

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 471 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GAATTCGGCA CGAGGCTCCC TCTCGTACTG CCATACTCCT GGGACGGGAT TCGGATAGGG
60
ATTTGCGGCG ATCCATTTCT CGATTCAAGG GGAAGAATCA TGGGGAAGTC CTACCCGACC
120
GTAAGCCAGG AGTACAAGAA GGCTGTCGAG AAATGCAAGA AGAAGTTGAG AGGCCTCATC
180
GCTGAGAAGA GCTGCGCTCC GCTCATGCTC CGCATCGCGT GGCACTCCGC CGGTACCTTC
240
GATGTGAAGA CGAAGACCGG AGGCCCGTTC GGGACCATGA AGCACGCCGC GGAGCTCAGC
300
CACGGGGCCA ACAGCGGGCT CGACGTTGCC GATCAGGTCT TGCAGCCGAT CAAGGATCAG
360
TTCCCCGTCA TCACTTATGC TGATTTCTAC CAGCTGGCTG GCGTCGTTGC TGTGGAAGTT
420
ACTGGTGGAC CTGAAGTTGC TTTTCACCCG GAAGAGAGGC AAACCACAAC C
471

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 487 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GAATTCGGCA CGAGCTCCCA CTTCTGTCTC GCCACCATTA CTAGCTTCAA AGCCCAGATC
60
TCAGTTTCGT GCTCTCTTCG TCATCTCTGC CTCTTGCCAT GGATCCGTAC AAGTATCGCC
120
CGTCCAGCGC TTACGATTCC AGCTTTTGGA CAACCAACTA CGGTGCTCCC GTCTGGAACA
180
ATGACTCATC GCTGACTGTT GGAACTAGAG GTCCGATTCT CCTGGAGGAC TACCATCTGA
240
TTGAGAAACT TGCCAACTTC GAGAGAGAGA GGATTCCTGA GCGGGTGGTC CATGCACGGG
300
GAGCCAGCGC GAAAGGGTTC TTCGAGGTCA CCCACGACAT CTCTCACTTG ACCTGTGCTG
360

ATTTCCTCCG GGCTCCTGGA GTCCAGACGC CCGTAATCGT CCGTTTCTCC ACCGTCATCC
420
ACGAGCGCGG CAGCCCGAAC CTCAGGGACC CTCGTGGTTT TGCAGTGAAG TTCTACACCA
480
GAGAGGG
487

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 684 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GAATTCCTGC AGCCCGGGGG ATCCACTAGT TCTAGAGCGG CCGCCACCGC GGTGGAGCTC
60
GCGCGCCTGC AGGTCGACAC TAGTGGATCC AAAGAATTCG GCACGAGGCC CGACGGCCAC
120
TTGTTGGACG CCATGGAAGC TCTCCGGAAA GCCGGGATTC TGGAACCGTT TAAACTGCAG
180
CCCAAGGAAG GACTGGCTCT CGTCAACGGC ACAGCGGTGG GATCCGCCGT GGCCGCGTCC
240
GTCTGTTTTG ACGCCAACGT GCTGGGCGTG CTGGCTGAGA TTCTGTCTGC GCTCTTCTGC
300
GAGGTGATGC AAGGGAAACC GGAGTTCGTA GATCCGTTAA CCCACCAGTT GAAGCACCAC
360
CCAGGGCAGA TCGAAGCCGC GGCCGTCATG GAGTTCCTCC TCGACGGTAG CGACTACGTG
420
AAAGAAGCAG CGCGGCTTCA CGAGAAAGAC CCGTTGAGCA AACCGAAACA AGACCGCTAC
480
GCTCTGCGAA CATCGCCACA GTGGTTGGGG CCTCCGATCG AAGTCATCCG CGCTGCTACT
540
CACTCCATCG AGCGGGAGAT CAATTCCGTC AACGACAATC CGTTAATCGA TGTCTCCAGG
600
GACATGGCTC TCCACGGCGG CAACTTCCAG GGAACACCCA TCGGAGTTTC CATGGACAAC
660
ATGCGAATCT CTTTGGCAGC CGTC
684

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 418 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GAATTCGGCA CGAGGACAAG GTCATAGGCC CTCTCTTCAA ATGCTTGGAT GGGTGGAAAG
60
GAACTCCTGG CCCATTCTGA AATAAATAAT CTTCCAAGAT CGCCTTTATA CAACGACTGC
120
TATGATTTGA GTCCTCGGAT CTTTTTGTTG ATGCAGTTGT TTACCGATCT GGAATTTGAT
180
TGGTCATAAA GCTTGATTTT GTTTTTCTTT CTTTTGTTTT ATACTGCTGG ATTTGCATCC
240
CATTGGATTT GCCAGAAATA TGTAAGGGTG GCAGATCATT TGGGTGATCT GAAACATGTA
300
AAAGTGGCGG ATCATTTGGG TAGCATGCAG ATCAGTTGGG TGATCGTGTA CTGCTTTCAC
360

TATTACTTAC ATATTTAAAG ATCGGGAATA AAAACATGAT TTTAATTGAA AAAAAAAA
418

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 479 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GATATCCCAA CGACCGAAAA CCTGTATTTT CAGGGCGCCA TGGGGATCCG GAATTCGGCA
60
CGAGCAAGGA AGAAAATATG GTTGCAGCAG CAGAAATTAC GCAGGCCAAT GAAGTTCAAG
120
TTAAAAGCAC TGGGCTGTGC ACGGACTTCG GCTCGTCTGG CAGCGATCCA CTGAACTGGG
180
TTCGAGCAGC CAAGGCCATG GAAGGAAGTC ACTTTGAAGA AGTGAAAGCG ATGGTGGATT
240
CGTATTTGGG AGCCAAGGAG ATTTCCATTG AAGGGAAATC TCTGACAATC TCAGACGTTG
300
CTGCCGTTGC TCGAAGATCG CAAGTGAAAG TGAAATTGGA TGCTGCGGCT GCCAAATCTA
360
GGGTCGAGGA GAGTTCAAAC TGGGTTCTCA CCCAGATGAC CAAGGGGACG GATACCTATG
420
GTGTCACTAC TGGTTTCGGA GCCACTTCTC ACAGGAGAAC GAACCAGGGA GCCGAGCTT
479

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1785 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TATCGATAAG CTTGATATCG AATTCCTGCA GCCCGGGGGA TCCACTAGTT CTAGAGCGGC
60
CGCCACCGCG GTGGAGCTCG CGCGCCTGCA GGTCGACACT AGTGGATCCA AAGAATTCGG
120
CACGAGGTTG CAGGTCGGGG ATGATTTGAA TCACAGAAAC CTCAGCGATT TTGCCAAGAA
180
ATATGGCAAA ATCTTTCTGC TCAAGATGGG CCAGAGGAAT CTTGTGGTAG TTTCATCTCC
240
CGATCTCGCC AAGGAGGTCC TGCACACCCA GGGCGTCGAG TTTGGGTCTC GAACCCGGAA
300
CGTGGTGTTC GATATCTTCA CGGGCAAGGG GCAGGACATG GTGTTCACCG TCTATGGAGA
360
TCACTGGAGA AAGATGCGCA GGATCATGAC TGTGCCTTTC TTTACGAATA AAGTTGTCCA
420
GCACTACAGA TTCGCGTGGG AAGACGAGAT CAGCCGCGTG GTCGCGGATG TGAAATCCCG
480
CGCCGAGTCT TCCACCTCGG GCATTGTCAT CCGTAGGCGC CTCCAGCTCA TGATGTATAA
540
TATTATGTAT AGGATGATGT TCGACAGGAG ATTCGAATCC GAGGACGACC CGCTTTTCCT
600
CAAGCTCAAG GCCCTCAACG GAGAGCGAAG TCGATTGGCC CAGAGCTTTG AGTACAATTA
660
TGGGGATTTC ATTCCCATTC TTAGGCCCTT CCTCAGAGGT TATCTCAGAA TCTGCAATGA
720

GATTAAAGAG AAACGGCTCT CTCTTTTCAA GGACTACTTC GTGGAAGAGC GCAAGAAGCT
780
CAACAGTACC AAGACTAGTA CCAACACCGG GGGAGCTCAA GTGTGCAATG GACCATATTT
840
TAGATGCTCA GGACAAGGGA GAGATCAATG AGGATAATGT TTTGTACATC GTTGAGAACA
900
TCAACGTTGC AGCAATTGAG ACAACGCTGT GGTCGATGGA ATGGGGAATA GCGGAGCTGG
960
TGAACCACCA GGACATTCAG AGCAAGGTGC GCGCAGAGCT GGACGCTGTT CTTGGACCAG
1020
GCGTGCAGAT AACGGAACCA GACACGACAA GGTTGCCCTA CCTTCAGGCG GTTGTGAAGG
1080
AAACCCTTCG TCTCCGCATG GCGATCCCGT TGCTCGTCCC CCACATGAAT CTCCACGACG
1140
CCAAGCTCGG GGGCTACGAT ATTCCGGCAG AGAGCAAGAT CCTGGTGAAC GCCTGGTGGT
1200
TGGCCAACAA CCCCGCCAAC TGGAAGAACC CCGAGGAGTT CCGCCCCGAG CGGTTCTTCG
1260
AGGAGGAGAA GCACACCGAA GCCAATGGCA ACGACTTCAA ATTCCTGCCT TCGGTGTGGG
1320
GAGGAGGAGC TGCCCGGGAA TCATTCTGGC GCTGCCTCTC CTCGCACTCT CCATCGGAAG
1380
ACTTGTTCAG AACTTCCACC TTCTGCCGCC GCCCGGGCAG AGCAAAGTGG ATGTCACTGA
1440
GAAGGGCGGG CAGTTCAGCC TTCACATTCT CAACCATTCT CTCATCGTCG CCAAGCCCAT
1500
AGCTTCTGCT TAATCCCAAC TTGTCAGTGA CTGGTATATA AATGCGCGCA CCTGAACAAA
1560
AAACACTCCA TCTATCATGA CTGTGTGTGC GTGTCCACTG TCGAGTCTAC TAAGAGCTCA
1620
TAGCACTTCA AAAGTTTGCT AGGATTTCAA TAACAGACAC CGTCAATTAT GTCATGTTTC
1680
AATAAAAGTT TGCATAAATT AAATGATATT TCAATATACT ATTTTGACTC TCCACCAATT
1740
GGGGAATTTT ACTGCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAA
1785

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 475 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GAATTCGGCA CGAGATTTCC ATGGACGATT CCGTTTGGCT TCAATTCGTT TCCTCTGGCT
60
GTCCTCGTCC TCGTTTTCCT TGTTCTTCCT CCGACTTTTT CTCTGGAAGC TATGGCGTAA
120
TAGGAACCTG CCGCCAGGAC CCCCGGCATG GCCGATCGTA GGGAACGTCC TTCAGATTGG
180
ATTTTCCAGC GGCGCGTTCG AGACCTCAGT GAAGAAATTC CATGAGAGAT ACGGTCCAAT
240
ATTCACTGTG TGGCTCGGTT CCCGCCCTCT GCTGATGATC ACCGACCGCG AGCTTGCCCA
300
CGAGGCGCTC GTACAGAAGG GCTCCGTCTT CGCTGACCGC CCGCCCGCCC TCGGGATGCA
360
GAAAATCTTC AGTAGCAACC AGCACAACAT CACTTCGGCT GAATACGGCC CGCTGTGGCG
420
GAGCCTTCGC AGGAATCTGG TTAAAGAAGC CCTGAGACTT CGGCGATGAA GGCTT
475

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 801 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GCTCCACCGA CGGTGGACGG TCCGCTACTC AGTAACTGAG TGGGATCCCC CGGGCTGACA
60
GGCAATTCGA TTTAGCTCAC TCATTAGGCA CCCCAGGCTT TACACTTTAT GCTTCCGGCT
120
CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTCACA CAGGAAACAG CTATGACCAT
180
GATTACGCCA AGCGCGCAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGA GCTCCACCGC
240
GGTGGCGGCC GCTCTAGAAC TAGTGGATCC AAAGAATTCG GCACGAGACC CAGTGACCTT
300
CAGGCCTGAG AGATTTCTTG AGGAAGATGT TGATATTAAG GGCCATGATT ACAGGCTACT
360
GCCATTCGGT GCAGGGCGCA GGATCTGCCC TGGTGCACAA TTGGGTATTA ATTTAGTTCA
420
GTCTATGTTG GGACACCTGC TTCATCATTT CGTATGGGCA CCTCCTGAGG GAATGAAGGC
480
AGAAGACATA GATCTCACAG AGAATCCAGG GCTTGTTACT TTCATGGCCA AGCCTGTGCA
540
GGCCATTGCT ATTCCTCGAT TGCCTGATCA TCTCTACAAG CGACAGCCAC TCAATTGATC
600
AATTGATCTG ATAGTAAGTT TGAATTTTGT TTTGATACAA AACGAAATAA CGTGCAGTTT
660
CTCCTTTTCC ATAGTCAACA TGCAGCTTTC TTTCTCTGAA GCGCATGCAG CTTTCTTTCT
720
CTGAAGCCCA ACTTCTAGCA AGCAATAACT GTATATTTTA GAACAAATAC CTATTCCTCA
780
AATTGAGTAT TTCTCTGTAG G
801

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 744 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GGGCCCCCCT TCGAGGTGGA CACTAGTGGA TCCAAAGAAT TCGGCACGAG GTTTTATCTG
60
AAGGACGCTG TGCTTGAAGG CTCCCAGCCA TTCACCAAAG CCCATGGAAT GAATGCGTTC
120
GAGTACCCGG CCATCGATCA GAGATTCAAC AAGATTTTCA ACAGGGCTAT GTCTGAGAAT
180
TCTACCATGT TGATGAACAA GATTTTGGAT ACTTACGAGG GTTTTAAGGA GGTTCAGGAG
240
TTGGTGGATG TGGGAGGAGG TATTGGGTCG ACTCTCAATC TCATAGTGTC TAGGTATCCC
300
CACATTTCAG GAATCAACTT CGACTTGTCC CATGTGCTGG CCGATGCTCC TCACTACCCA
360
GCTGTGAAAC ATGTGGGTGG AGACATGTTT GATAGTGTAC CAAGTGGCCA AGCTATTTTT
420
ATGAAGTGGA TTCTGCATGA TTGGAGCGAT GATCATTGCA GGAAGCTTTT GAAGAATTGT
480

AGCCCTTGCA TTGCCCGATG ATGGAAAGAT TCTAGCCATG GACATCAACA GAGAGAACTA
540
TGATATCGGA TTGCCTATAA TT
562

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1074 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

TCGTGCCGCT CGATCCTCAC AGGCCCTTTT TATTTCCCTG GTGAACGATA CGATGGGCTC
60
GCACGCTGAG AATGGCAACG GGGTGGAGGT TGTTGATCCA ACGGACTTAA CTGACATCGA
120
GAATGGGAAA CCAGGTTATG ACAAGCGTAC GCTGCCTGCG GACTGGAAGT TTGGAGTGAA
180
GCTTCAAAAC GTTATGGAAG AATCCATTTA CAAGTACATG CTGGAAACAT TCACCCGCCA
240
TCGAGAGGAC GAGGCGTCCA AGGAGCTCTG GGAACGAACA TGGAACCTGA CACAGAGAGG
300
GGAGATGATG ACATTGCCAG ATCAGGTGCA GTTCCTGCGC TTGATGGTAA AGATGTCAGG
360
TGCTAAAAAG GCATTGGAGA TCGGAGTTTT CACTGGCTAT TCATTGCTCA ATATCGCTCT
420
CGCTCTTCCT TCTGATGGCA AGGTGGTAGC TGTGGATCCA GGAGATGACC CCAAATTTGG
480
CTGGCCCTGC TTCGTTAAGG CTGGAGTTGC AGACAAAGTG GAGATCAAGA AAACTACAGG
540
GTTGGACTAT TTGGATTCCC TTATTCAAAA GGGGGAGAAG GATTGCTTCG ACTTTGCATT
600
CGTGGACGCA GACAAAGTGA ACTACGTGAA CTATCATCCA CGGCTGATGA AGTTAGTGCG
660
CGTGGGGGGC GTCATAATTT ACGACGACAC CCTCTGGTTT GGTCTGGTGG GAGGAAAGGA
720
TCCCCACAAC CTGCTTAAGA ATGATTACAT GAGGACTTCT CTGGAGGGTA TCAAGGCCAT
780
CAACTCCATG GTAGCCAACG ACCCCAACTT GGAGGTCGCC ACAGTCTTTA TGGGATATGG
840
TGTCACTGTT TGTTACCGCA CTGCTTAGTT AGCTAGTCCT CCGTCATTCT GCTATGTATG
900
TATATGATAA TGGCGTCGAT TTCTGATATA GGTGGTTTTT CAATGTTTCT ATCGTCATGT
960
TTTCTGTTTA GCCAGAATGT TTCGATCGTC ATGGTTTCTG TTAAAGCCAG AATAAAATTA
1020
GCCGCTTGCA GTTCAAAAAA AAAAAAAAAA AAAAACTCGA GACTAGTTCT CTTC
1074

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1075 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TCGGAGCTCT CGAATCCTCA CAGGCCCTTT TTATTTCCCT GGTGAACGAT ACGATGGGCT
60

CACAAGGCGT TGCCAGAGAA GGGGAAGGTG ATTGCGGTGG ACACCATTCT CCCAGTGGCT
540
GCAGAGACAT CTCCTTATGC TCGTCAGGGA TTTCATACAG ATTTACTGAT GTTGGCATAC
600
AACCCAGGGG GCAAGGAACG CACAGAGCAA GAATTTCAAG ATTTAGCTAA GGAGACGGGA
660
TTTGCAGGTG GTGTTGAACC TGTATGTTGT GTCAATGGAA TGTGGGTAAT GGAATTCCTG
720
CAGCCCGGGG GATCCACTAG TTCT
744

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 426 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTGGCCCTGG AAGTAGTGTG CGCGACATGG ATTCCTTGAA TTTGAACGAG TTTATGTTGT
60
GGTTTCTCTC TTGGCTTGCT CTCTACATTG GATTTCGTTA TGTTTTGAGA TCGAACTTGA
120
AGCTCAAGAA GAGGCGCCTC CCGCCGGGCC CATCGGGATG GCCAGTGGTG GGAAGTCTGC
180
CATTGCTGGG AGCGATGCCT CACGTTACTC TCTACAACAT GTATAAGAAA TATGGCCCCG
240
TTGTCTATCT CAAACTGGGG ACGTCCGACA TGGTTGTGGC CTCCACGCCC GCTGCAGCTA
300
AGGCGTTTCT GAAGACTTTG GATATAAACT TCTCCAACCG GCCGGGAAAT GCAGGAGCCA
360
CGTACATCGC CTACGATTCT CAGGACATGG TGTGGGCAGC GTATGGAGGA CGGTGGAAGA
420
TGGAGC
426

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 562 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CAGTTCGAAA TTAACCTCAC TAAAGGGAAC AAAAGCTGGA GTTCGCGCGC CTGCAGGTCG
60
ACACTAGTGG ATCCAAAGAA TTCGGCACGA GCTTTGAGGC AACCTACATT CATTGAATCC
120
CAGGATTTCT TCTTGTCCAA ACAGGTTTAA GGAAATGGCA GGCACAAGTG TTGCTGCAGC
180
AGAGGTGAAG GCTCAGACAA CCCAAGCAGA GGAGCCGGTT AAGGTTGTCC GCCATCAAGA
240
AGTGGGACAC AAAAGTCTTT TGCAGAGCGA TGCCCTCTAT CAGTATATAT TGGAAACGAG
300
CGTGTACCCT CGTGAGCCCG AGCCAATGAA GGAGCTCCGC GAAGTGACTG CCAAGCATCC
360
CTGGAACCTC ATGACTACTT CTGCCGATGA GGGTCAATTT CTGGGCCTCC TGCTGAAGCT
420
CATTAACGCC AAGAACACCA TGGAGATTGG GGTGTACACT GGTTACTCGC TTCTCAGCAC
480

AAGCCGACGA AACCCAATGC CCGGCCGTGA CAATCCACCC GGACGATGTC GTGGCGTTGC
660
CCTATTCTTC CGGAACCACG GGGCTCCCCA AGGGCGTGAT GTTAACGCAC AAAGGCCTGG
720
TGTCCAGCGT TGCCCAGCAG GTCGATGGTG AAAATCCCAA TCTGTATTTC CATTCCGATG
780
ACGTGATACT CTGTGTCTTG CCTCTTTTCC ACATCTATTC TCTCAATTCG GTTCTCCTCT
840
GCGCGCTCAG AGCCGGGGCT GCGACCCTGA TTATGCAGAA ATTCAACCTC ACGACCTGTC
900
TGGAGCTGAT TCAGAAATAC AAGGTTACCG TTGCCCCAAT TGTGCCTCCA ATTGTCCTGG
960
ACATCACAAA GAGCCCCATC GTTTCCCAGT ACGATGTCTC GGCCGTCCGG ATAATCATGT
1020
CCGGCGCTGC GCCTCTCGGG AAGGAACTCG AAGATGCCCT CAGAGAGCGT TTTCCCAAGG
1080
CCATTTTCGG GCAGGGCTAC GGCATGACAG AAGCAGGCCC GGTGCTGGCA ATGAACCTAG
1140
CCTTCGCAAA GAATCCTTTC CCCGTCAAAT CTGGCTCCTG CGGAACAGTC GTCCGGAACG
1200
CTCAAATAAA GATCCTCGAT ACAGAAACTG GCGAGTCTCT CCCGCACAAT CAAGCCGGCG
1260
AAATCTGCAT CCGCGGACCC GAAATAATGA AAGGATATAT TAACGACCCG GAATCCACGG
1320
CCGCTACAAT CGATGAAGAA GGCTGGCTCC ACACAGGCGA CGTCGGGTAC ATTGACGATG
1380
ACGAAGAAAT CTTCATAGTC GACAGAGTAA AGGAGATTAT CAAATATAAG GGCTTCCAGG
1440
TGGCTCCTGC TGAGCTGGAA GCTTTACTTG TTGCTCATCC GTCAATCGCT GACGCAGCAG
1500
TCGTTCCTCA AAAGCACGAG GAGGCGGGCG AGGTTCCGGT GGCGTTCGTG GTGAAGTCGT
1560
CGGAAATCAG CGAGCAGGAA ATCAAGGAAT TCGTGGCAAA GCAGGTGATT TTCTACAAGA
1620
AAATACACAG AGTTTACTTT GTGGATGCGA TTCCTAAGTC GCCGTCCGGC AAGATTCTGA
1680
GAAAGGATTT GAGAAGCAGA CTGGCAGCAA AATGAAAATG AATTTCCATA TGATTCTAAG
1740
ATTCCTTTGC CGATAATTAT AGGATTCCTT TCTGTTCACT TCTATTTATA TAATAAAGTG
1800
GTGCAGAGTA AGCGCCCTAT AAGGAGAGAG AGAGCTTATC AATTGTATCA TATGGATTGT
1860
CAACGCCCTA CACTCTTGCG ATCGCTTTCA ATATGCATAT TACTATAAAC GATATATGTT
1920
TTTTTTATAA ATTTACTGCA CTTCTCGTTC AAAAAAAAAA A
1961

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1010 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACAAACTTG GTCGTTTGTT TAGGTTTTGC TGCAGGTGAA CACTAATATG GAAGGCCAGA
60
TTGCAGCATT AAGCAAAGAA GATGAGTTCA TTTTTCACAG CCCTTTTCCT GCAGTACCTG
120
TTCCAGAGAA TATAAGTCTT TTCCAGTTTG TTCTGGAAGG TGCTGAGAAA TACCGTGATA
180
AGGTGGCCCT CGTGGAGGCC TCCACAGGGA AGGAGTACAA CTATGGTCAG GTGATTTCGC
240

CGCACGCTGA GAATGGCAAC GGGGTGGAGG TTGTTGATCC AACGGACTTA ACTGACATCG
120
AAGAATGGGA AACCAGGTTA TGACAAGCGT CGCTGCCTGC GGACTGGAAG TTTGGAGTGA
180
AGCTTCAAAA CGTTATGGAA GAATCCATTT ACAAGTACAT GCTGGAAACA TTCACCCGCC
240
ATCGAGAGGA CGAGGCGTCC AAGGAGCTCT GGGAACGAAC ATGGAACCTG ACACAGAGAG
300
GGGAGATGAT GACATTGCCA GATCAGGTGC AGTTCCTGCG CTTGATGGTA AAGATGTCAG
360
GTGCTAAAAA GGCATTGGAG ATCGGAGTTT TCACTGGCTA TTCATTGCTC AATATCGCTC
420
TCGCTCTTCC TTCTGATGGC AAGGTGGTAG CTGTGGATCC AGGAGATGAC CCCAAATTTG
480
GCTGGCCCTG CTTCGTTAAG GCTGGAGTTG CAGACAAAGT GGAGATCAAG AAAACTACAG
540
GGTTGGACTA TTTGGATTCC CTTATTCAAA AGGGGGAGAA GGATTGCTTC GACTTTGCAT
600
TCGTGGACGC AGACAAAGTG AACTACGTGA ACTATCATCC ACGGCTGATG AAGTTAGTGC
660
GCGTGGGGGG CGTCATAATT TACGACGACA CCCTCTGGTT TGGTCTGGTG GGAGGAAAGG
720
ATCCCCACAA CCTGCTTAAG AATGATTACA TGAGGACTTC TCTGGAGGGT ATCAAGGCCA
780
TCAACTCCAT GGTAGCCAAC GACCCCAACT TGGAGGTCGC CACAGTCTTT ATGGGATATG
840
GTGTCACTGT TTGTTACCGC ACTGCTTAGT TAGCTAGTCC TCCGTCATTC TGCTATGTAT
900
GTATATGATA ATGGCGTCGA TTTCTGATAT AGGTGGTTTT TCAATGTTTC TATCGTCATG
960
TTTTCTGTTT AGCCAGAATG TTTCGATCGT CATGGTTTCT GTTAAAGCCA GAATAAAATT
1020
AGCCGCTTGC AGTTCAAAAA AAAAAAAAAA AAAAAACTCG AGACTAGTTC TCTTC
1075

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1961 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GTTTTCCGCC ATTTTTCGCC TGTTTCTGCG GAGAATTTGA TCAGGTTCGG ATTGGGATTG
60
AATCAATTGA AAGGTTTTTA TTTTCAGTAT TTCGATCGCC ATGGCCAACG GAATCAAGAA
120
GGTCGAGCAT CTGTACAGAT CGAAGCTTCC CGATATCGAG ATCTCCGACC ATCTGCCTCT
180
TCATTCGTAT TGCTTTGAGA GAGTAGCGGA ATTCGCAGAC AGACCCTGTC TGATCGATGG
240
GGCGACAGAC AGAACTTATT GCTTTTCAGA GGTGGAACTG ATTTCTCGCA AGGTCGCTGC
300
CGGTCTGGCG AAGCTCGGGT TGCAGCAGGG GCAGGTTGTC ATGCTTCTCC TTCCGAATTG
360
CATCGAATTT GCGTTTGTGT TCATGGGGGC CTCTGTCCGG GGCGCCATTG TGACCACGGC
420
CAATCCTTTC TACAAGCCGG GCGAGATCGC CAAACAGGCC AAGGCCGCGG GCGCGCGCGA
480
TCATAGTTAC CCTGGCAGCT TATGTGGAGA AACTGGCCGA TCTGCAGAGC CACGATGTGC
540
TCGTCATCAC AATCGATGAT GCTCCCAAGG AAGGTTGCCA ACATATTTCC GTTCTGACCG
600

TCACAAGGAA TGTTGCAGCT GGGCTCGTGG ACAAAGGCAT TCAAAAGGGC GATGTTGTAT
300
TTGTTCTGCT TCCAAATATG GCAGAATACC CCATTATTGT GCTGGGAATA ATGTTGGCCG
360
GCGCAGTGTT TTCTGGGGCA AATCCTTCTG CACACATCAA TGAAGTTGAA AAACATATCC
420
AGGATTCTGG AGCAAAGATT GTTGTGACAG TTGGGTCTGC TTATGAGAAG GTGAGGCAAG
480
TGAAACTGCC TGTTATTATT GCAGATAACG AGCATGTCAT GAACACAATT CCATTGCAGG
540
AAATTTTTGA GAGAAACTAT GAGGCCGCAG GGCCTTTTGT ACAAATTTGT CAGGATGATC
600
TGTGTGCACT CCCTTATTCC TCTGGCACCA CAGGGGCCTC TAAAGGTGTC ATGCTCACTC
660
ACAGAAATCT GATTGCAAAT CTGTGCTCTA GCTTGTTTGA TGTCCATGAA TCTCTTGTAG
720
GAAATTTCAC CACGTTGGGG CTGATGCCAT TCTTTCACAT ATATGGCATC ACGGGCATCT
780
GTTGCGCCAC TCTTCGCAAC GGAGGCAAGG TCGTGGTCAT GTCCAGATTC GATCTCCGAC
840
ACTTTATCAG TTCTTTGATT ACTTATGAGG TCAACTTCGC GCCTATTGTC CCGCCTATAA
900
TGCTCTCCCT CCGGTTTAAA AATCCTATCG TTAACGAGTT CGATCTCAGC CGCTTGAAAC
960
TCCAAAGCTG TTCATGACTG CGGCTGCTCC ACTGGCGCCG GATCTACTGC
1010

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 741 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GAATTCGGCA CGAGACCATT TCCAGCTAAT ATTGGCATAG CAATTGGTCA TTCTATCTTT
60
GTCAAAGGAG ATCAAACAAA TTTTGAAATT GGACCTAATG GTGTGGAGGC TAGTCAGCTA
120
TACCCAGATG TGAAATATAC CACTGTCGAT GAGTACCTCA GCAAATTTGT GTGAAGTATG
180
CGAGATTCTC TTCCACATGC TTCAGAGATA CATAACAGTT TCAATCAATG TTTGTCCTAG
240
GCATTTGCCA AATTGTGGGT TATAATCCTT CGTAGGTGTT TGGCAGAACA GAACCTCCTG
300
TTTAGTATAG TATGACGAGC TAGGCACTGC AGATCCTTCA CACTTTTCTC TTCCATAAGA
360
AACAAATACT CACCTGTGGT TTGTTTTCTT TCTTTCTGGA ACTTTGGTAT GGCAATAATG
420
TCTTTGGAAA CCGCTTAGTG TGGAATGCTA AGTACTAGTG TCCAGAGTTC TAAGGGAGTT
480
CCAAAATCAT GGCTGATGTG AACTGGTTGT TCCAGAGGGT GTTTACAACC AACAGTTGTT
540
CAGTGAATAA TTTTGTTAGA GTGTTTAGAT CCATCTTTAC AAGGCTATTG AGTAAGGTTG
600
GTGTTAGTGA ACGGAATGAT GTCAAATCTT GATGGGCTGA CTGACTCTCT TGTGATGTCA
660
AATCTTGATG GATTGTGTCT TTTTCAATGG TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
720
AAAAAAAAAA AAAAAAAAAA A
741

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 643 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

CTCATCTCGG AGTTGCAGGC TGCAGCTTTT GGCCCAAAGC ATGATATCAG ATCAAACGAC
60
GCAGATGAAG CAAACGGATC AAACAGTTTG CGTTACTGGA GCAGCGGGTT TCATTGCCTC
120
ATGGCTTGTC AAGATGCTCC TCATCAGAGG TTACACTGTC AGAGCAGCAG TTCGGACCAA
180
CCCAGCTGAT GATAGGTGGA AGTATGAGCA TCTGCGAGAG TTGGAAGGAG CAAAAGAGAG
240
GCTTGAGCTT GTGAAAGCTG ATATTCTCCA TTACCAGAGC TTACTCACAG TCATCAGAGG
300
TTGCCACGGT GTCTTTCACA TGGCTTCAGT TCTCAATGAT GACCCTGAGC AAGTGATAGA
360
ACCAGCAGTC GAAGGGACGA GGAATGTGAT GGAGGCCTGC GCAGAAACTG GGGTGAAGCG
420
CGTTGTTTTT ACTTCTTCCA TCGGCGCAGT TTACATGAAT CCTCATAGAG ACCCGCTCGC
480
GATTGTCCAT GATGACTGCT GGAGCGATTT GACTACTGCG TACAAACCAA GAATTGGTAT
540
TGCTATGCAA AAACCTTGGC AGAGAAATCT GCATGGGATA TTGCTAAGGG AAGGAATTTA
600
GAGCTTGCAG TGATAAATCC AGGCCTGGCC TTAGGTCCCT TGA
643

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 441 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GAATTCGGCA CGAGAATTTT TCTGTGGTAA GCATATCTAT GGCTCAAACC AGAGAGAAGG
60
ACGATGTCAG CATAACAAAC TCCAAAGGAT TGGTATGCGT GACAGGAGCG GCTGGTTACT
120
TGGCATCTTG GCTTATCAAG CGTCTCCTCC AGTGTGGTTA CCAAGTGAGA GGAACTGTGC
180
GGGATCCTGG CAATGAGAAA AAGATGGCTC ATTTATGGAA GTTAGATGGG GCGAAAGAGA
240
GACTGCAACT AATGAAAGCT GATTTAATGG ACGAGGGCAG CTTCGATGAG GTCATCAGAG
300
GCTGCCATGG TGTTTTTCAC ACAGCGTCTC CAGTCGTGGG TGTCAAATCA GATCCCAAGA
360
TATGGTATGC TCTGGCCAAG ACTTTAGCAG AAAAAGCAGC ATGGGATTTT GCCCAAGAAA
420
ACCATCTGGA CATGGTTGCA G
441

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 913 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GAATTCGGCA CGAGGAAAAC ATCATCCAGG CATTTTGGAA ATTTAGCTCG CCGGTTGATT
60
CAGGATCCTG CAATGGCTTT TGGCGAAGAG CAGACTGCCT TGCCACAAGA AACGCCTTTG
120
AATCCTCCGG TCCATCGAGG AACAGTGTGC GTTACAGGAG CTGCTGGGTT CATAGGGTCA
180
TGGCTCATCA TGCGATTGCT TGAGCGAGGA TATAGTGTTA GAGCAACTGT GCGAGACACT
240
GGTAATCCTG TAAAGACAAA GCATCTGTTG GATCTGCCGG GGGCAAATGA GAGATTGACT
300
CTCTGGAAAG CAGATTTGGA TGATGAAGGA AGCTTTGATG CTGCCATTGA TGGGTGTGAG
360
GGTGTTTTCC ATGTTGCCAC TCCCATGGAT TTCGAGTCCG AGGATCCCGA GAATGAGATA
420
ATTAAGCCAA CAATCAACGG GGTCTTGAAT GTTATGAGAT CGTGTGCAAA AGCCAAGTCC
480
GTGAAGCGAG TTGTTTTCAC GTCATCTGCT GGGACTGTGA ATTTTACAGA TGATTTCCAA
540
ACACCAGGCA AAGTTTTTGA CGAATCATGC TGGACCAACG TGGATCTTTG CAGAAAAGTT
600
AAAATGACAG GATGGATGTA CTTTGTATCG AAGACATTAG CAGAGAAAGC TGCTTGGGAT
660
TTTGCAGAGG AGAACAAGAT CGATCTCATT ACTGTTATCC CCACATTGGT CGTTGGACCA
720
TTCATTATGC AGACCATGCC ACCGAGCATG ATCACAGCCT TGGCACTGTT AACGCGGAAT
780
GAACCCCACT ACATGATACT GAGACAGGTA CAGCTGGTTC ACTTGGATGA TCTCTGTATG
840
TCACATATCT TTGTATATGA ACATCCTGAA GCAAAGGGCA GATACATCTC TTCCACATGT
900
GATGCTACCC ATT
913

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 680 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GAATTCGGCA CGAGATCAAT TTTTGCATAT TATTAAAAAG TAAGTGTATT CGTTCTCTAT
60
ATTGATCAGT CACAGAGTCA TGGCCAGTTG TGGTTCCGAG AAAGTAAGAG GGTTGAATGG
120
AGATGAAGCA TGCGAAGAGA ACAAGAGAGT GGTTTGTGTA ACTGGGGCAA ATGGGTACAT
180
CGGCTCTTGG CTGGTCATGA GATTACTGGA ACATGGCTAT TATGTTCATG GAACTGTTAG
240
GGACCCAGAA GACACAGGGA AGGTTGGGCA TTTGCTGCGG CTCCCAGGGG CAAGTGAGAA
300
GCTAAAGCTG TTCAAGGCAG AGCTTAACGA CGAAATGGCC TTTGATGATG CTGTGAGCGG
360
TTGTCAAGGG GTTTTCCACG TTGCCAAGCC TGTTAATCTG GACTCAAACG CTCTTCAGGG
420
GGAGGTTGTT GGTCCTGCGG TGAGGGGAAC AGTAAATCTG CTTCGAGCCT GCGAACGATC
480

GGGCACTGTG AAACGAGTGA TACATACCTC GTCCGTTTCA GCAGTGAGAT TCACTGGGAA
540
ACCTGACCCC CCTGATACTG TGCTGGATGA ATCTCATTGG ACTTCGGTCG AGTATTGCAG
600
AAAGACAAAG ATGGTCGGAT GGATGTACTA CATCGCCAAC ACTTATGCAG AAGAGGGAGC
660
CCATAAGTTC GGATCAGAGA
680

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 492 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GAATTCGGCA CGAGGCTGGT TCAAGTGTCA GCCCAATGGC CTCCCCTACA GAGAATCCCC
60
AGATTTCAGA AGAGCTGCTA AATCATGAGA TCCATCAAGG AAGTACAGTA TGTGTGACAG
120
GAGCTGCTGG CTTCATAGGA TCATGGCTCG TCATGCGTTT GCTTGAGCGA GGATATACTG
180
TTAGAGGAAC TGTGCGAGAC ACTGGTAATC CGGTGAAGAC GAAGCATCTA TTGGATCTGC
240
CTGGGGCGAA TGAGAGGTTA ACTCTCTGGA AAGCAGATTT GGATGATGAA GGAAGCTTTG
300
ACGCCGCCAT TGATGGTTGT GAGGGAGTTT TCCATGTTGC CACTCCCATG GATTTTGAAT
360
CCGAGGACCC CGAGAACGAG ATAATTAAAC CCGCTGTCAA TGGGATGTTG AATGTTTTGA
420
GATCGTGTGG GAAAACCAAG TCTATGAAGC GAGTTGTTTT CACGTCGTCT GCTGGGACTC
480
TGCTTTTTAC GG
492

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 524 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GAATTCGGCA CGAGCTTGTT CAAAGTCACA TATCTTATTT TCTTTGTGAT ATCTGCAATT
60
TCCAAGCTTT TCGTCTACCT CCCTGAAAAG ATGAGCGAGG TATGCGTGAC AGGAGGCACA
120
GGCTTCATAG CTGCTTATCT CATTCGTAGT CTTCTCCAGA AAGGTTACAG AGTTCGCACT
180
ACAGTTCGCA ACCCAGATAA TGTGGAGAAG TTTAGTTATC TGTGGGATCT GCCTGGTGCA
240
AACGAAAGAC TCAACATCGT GAGAGCAGAT TTGCTAGAGG AAGGCAGTTT TGATGCAGCA
300
GTAGATGGTG TAGATGGAGT ATTCCATACT GCATCACCTG TCTTAGTCCC ATATAACGAG
360
CGCTTGAAGG AAACCCTAAT AGATCCTTGT GTGAAGGGCA CTATCAATGT CCTCAGGTCC
420
TGTTCAAGAT CACCTTCAGT AAAGCGGGTG GTGCTTACAT CCTCCTGCTC ATCAATACCG
480

ATACGACTAT AATAGCTTAG AGCGTTCCCT GCTGGACTGA GTCA
524

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 417 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

TCCTAATTGT TCGATCCTCC CTTTTAAAGC CCTTCCCTGG CCTTCATTCC AGGTCACAGA
60
GTTGTTCATG CAGTGCTAGC AGGAGGAGCA GCGTTGCAAT TGGGGAAAAT TCCAAAATCA
120
ATAACGAGAG GACAGAAGTA AGTTTGTGGA AATAGCAACC ATGCCGGTGT TTCCTTCTGG
180
TCTGGACCCC TCTGAGGACA ATGGCAAGCT CGTTTGTGTC ATGGATGCGT CCAGTTATGT
240
AGGTTTGTGG ATTGTTCAGG GCCTTCTTCA ACGAGGCTAT TCAGTGCATG CCACGGTGCA
300
GAGAGACGCT GGCGAGGTTG AGTCTCTCAG AAAATTGCAT GGGGATCGAT TGCAGATCTT
360
CTATGCAGAT GTCTTGGATT ATCACAGCAT TACTGATGCG CTCAAGGGCT GTTCTGG
417

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 511 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATGACACGAA TTTGTGCCTC TCTCTGACCA GAGCTTGAAG CTCTGTCTTC TCTGATATCG
60
CTTCATTCCA TCATCCAGGA GCTTCTGTTA TATCCATTTC CTCAAAATGG ATGCCTACCT
120
TGAAGAAAAT GGATACGGCG CTTCCAATTC TCGGAAATTA ATGTGCCTTA CCGGGGGCTG
180
GAGTTTCCTG GGGATTCATA TCGCAAGAAT GCTGCTCGGC CGGGGTTACT CAGTCCGTTT
240
CGCAATTCCG GTAACGCCAG AAGAGGCAGG CTCACTTATG GAATCCGAAG AAGCATTATC
300
GGGGAAGCTG GAGATATGCC AAGCCGATCT CTTGGATTAT CGCAGCGTTT TCGGCAACAT
360
CAATGGTTGC TCCGGAGTCT TCCACGTCCC TGCGCCCTGT GATCATCTGG ATGGATTACA
420
GGAGTATCCG GTATGATTAG TTTAATAGAT TGACGGGGTA TCCTGTATGA ATTAGTTTAT
480
GAATTTAAGG TTTTCTTAGA ATTTGGATAC T
511

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 609 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CATTGATAGT TGATGGAAGA CCATCAGTAA AGCATGAAAA AGAAATTGTT CCAAGGTGAA
60
GAAGTCAGTT GCTCCAGCAG AACCTTTTTA GCAATTGTTT TTGTATCCTT TTTGCCTTTG
120
AATATGTAAT CCATAAACTT ATGCAGGAAG TGCCTCGTGC CGAATTCGGC ACGAGAATCA
180
CTGACCTTCA CATATTTATT CCAATTCTAA TATCTCTACT CGCTGTCTAC CTGATTTTTC
240
AGTGGCGAAC CAACTTGACA GGGTTGGACA TGGCCAACAG CAGCAAGATT CTGATTATTG
300
GAGGAACAGG CTACATTGGT CGTCATATAA CCAAAGCCAG CCTTGCTCTT GGTCATCCCA
360
CATTCCTTCT TGTCAGAGAG ACCTCCGCTT CTAATCCTGA GAAGGCTAAG CTTCTGGAAT
420
CCTTCAAGGC CTCAGGTGCT ATTATACTCC ATGGATCTTT GGAGGACCAT GCAAGTCTTG
480
TGGAGGCAAT CAAGAAAGTT GATGTAGTTA TCTCGGCTGT CAAGGGACCA CAGCTGACGG
540
TTCAAACAGG ATATTTATCC AGGGTATTTA AAGGGAGGGT TGGAACCCAT CAAGAAGGGT
600
TTTGGCCAA
609

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 474 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GCAAGATAGG TTTTATTCTT CTGGAGTTGG GTGAGGCTTG GAAATTTAAG TAAAAAGGGT
60
GCATAGCAAT TAAGCAGTTG CAGCCATGGC GGTCTGTGGA ACTGAAGTAG CTCATACTGT
120
GCTCTATGTA GCTGCAGACA TGGTGGAAAA CAACACGTCT ATTGTGACCA CCTCTATGGC
180
TGCAGCAAAT TGTGAGATGG AGAAGCCTCT TCTAAATTCC TCTGCCACCT CAAGAATACT
240
GGTGATGGGA GCCACAGGTT ACATTGGCCG TTTTGTTGCC CAAGAAGCTG TTGCTGCTGG
300
TCATCCTACC TATGCTCTTA TACGCCCGTT TGCTGCTTGT GACCTGGCCA AAGCACAGCG
360
CGTCCAACAA TTGAAGGATG CCGGGGTCCA TATCCTTTAT GGGTCTTTGA GTGATCACAA
420
CCTCTTAGTA AATACATTGA AGGACATGGG CCGTTGTTAT CTCTACCATT GGAG
474

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 474 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

GCAAGATAGG TTTTATTCTT CTGGAGTTGG GTGAGGCTTG GAAATTTAAG TAAAAAGGGT
60
GCATAGCAAT TAAGCAGTTG CAGCCATGGC GGTCTGTGGA ACTGAAGTAG CTCATACTGT
120
GCTCTATGTA GCTGCAGACA TGGTGGAAAA CAACACGTCT ATTGTGACCA CCTCTATGGC
180
TGCAGCAAAT TGTGAGATGG AGAAGCCTCT TCTAAATTCC TCTGCCACCT CAAGAATACT
240
GGTGATGGGA GCCACAGGTT ACATTGGCCG TTTTGTTGCC CAAGAAGCTG TTGCTGCTGG
300
TCATCCTACC TATGCTCTTA TACGCCCGTT TGCTGCTTGT GACCTGGCCA AAGCACAGCG
360
CGTCCAACAA TTGAAGGATG CCGGGGTCCA TATCCTTTAT GGGTCTTTGA GTGATCACAA
420
CCTCTTAGTA AATACATTGA AGGACATGGG CCGTTGTTAT CTCTACCATT GGAG
474

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 608 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CATTGATAGT TGATGGAAGA CCATCAGTAA AGCATGAAAA AGAAATTGTT CCAAGGTGAA
60
GAAGTCAGTT GCTCCAGCAG AACCTTTTTA GCAATTGTTT TTGTATCCTT TTTGCCTTTG
120
AATATGTAAT CCATAAACTT ATGCAGGAAG TGCCTCGTGC CGAATTCGGC ACGAGAATCA
180
CTGACCTTCA AATATTTATT CCAATTCTAA TATCTCTACT CGCTGTCTAC CTGATTTTTC
240
AGTGGCGAAC CAACTTGACA GGGTTGGACA TGGCCAACAG CAGCAAGATT CTGATTATTG
300
GAGGAACAGG CTACATTGGT CGTCATATAA CCAAAGCCAG CCTTGCTCTT GGTCATCCCA
360
CATTCCTTCT TGTCAGAGAG ACCTCCGCTT CTAATCCTGA GAAGGCTAAG CTTCTGGAAT
420
CCTTCAAGGC CTCAGGTGCT ATTATACTCC ATGGATCTTT GGAGGACCAT GCAAGTCTTG
480
TGGAGGCAAT CAAGAAAGTT GATGTAGTTA TCTCGGCTGT CAAGGGACCA CAGCTGACGG
540
ATCAAACAGG ATATTTATCC AGGGTATTTA AAGGGAGGTT GGAACCCATC AAGAAGGGTT
600
TTGGCCAA
608

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1474 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GAATTCGGCA CGAGAAAACG TCCATAGCTT CCTTGCCAAC TGCAAGCAAT ACAGTACAAG
60
AGCCAGACGA TCGAATCCTG TGAAGTGGTT CTGAAGTGAT GGGAAGCTTG GAATCTGAAA
120

AAACTGTTAC AGGATATGCA GCTCGGGACT CCAGTGGCCA CTTGTCCCCT TACACTTACA
180
ATCTCAGAAA GAAAGGACCT GAGGATGTAA TTGTAAAGGT CATTTACTGC GGAATCTGCC
240
ACTCTGATTT AGTTCAAATG CGTAATGAAA TGGACATGTC TCATTACCCA ATGGTCCCTG
300
GGCATGAAGT GGTGGGGATT GTAACAGAGA TTGGCAGCGA GGTGAAGAAA TTCAAAGTGG
360
GAGAGCATGT AGGGGTTGGT TGCATTGTTG GGTCCTGTCG CAGTTGCGGT AATTGCAATC
420
AGAGCATGGA ACAATACTGC AGCAAGAGGA TTTGGACCTA CAATGATGTG AACCATGACG
480
GCACACCTAC TCAGGGCGGA TTTGCAAGCA GTATGGTGGT TGATCAGATG TTTGTGGTTC
540
GAATCCCGGA GAATCTTCCT CTGGAACAAG CGGCCCCTCT GTTATGTGCA GGGGTTACAG
600
TTTTCAGCCC AATGAAGCAT TTCGCCATGA CAGAGCCCGG GAAGAAATGT GGGATTTTGG
660
GTTTAGGAGG CGTGGGGCAC ATGGGTGTCA AGATTGCCAA AGCCTTTGGA CTCCACGTGA
720
CGGTTATCAG TTCGTCTGAT AAAAAGAAAG AAGAAGCCAT GGAAGTCCTC GGCGCCGATG
780
CTTATCTTGT TAGCAAGGAT ACTGAAAAGA TGATGGAAGC AGCAGAGAGC CTAGATTACA
840
TAATGGACAC CATTCCAGTT GCTCATCCTC TGGAACCATA TCTTGCCCTT CTGAAGACAA
900
ATGGAAAGCT AGTGATGCTG GGCGTTGTTC CAGAGCCGTT GCACTTCGTG ACTCCTCTCT
960
TAATACTTGG GAGAAGGAGC ATAGCTGGAA GTTTCATTGG CAGCATGGAG GAAACACAGG
1020
AAACTCTAGA TTTCTGTGCA GAGAAGAAGG TATCATCGAT GATTGAGGTT GTGGGCCTGG
1080
ACTACATCAA CACGGCCATG GAAAGGTTGG AGAAGAACGA TGTCCGTTAC AGATTTGTGG
1140
TGGATGTTGC TAGAAGCAAG TTGGATAATT AGTCTGCAAT CAATCAATCA GATCAATGCC
1200
TGCATGCAAG ATGAATAGAT CTGGACTAGT AGCTTAACAT GAAAGGGAAA TTAAATTTTT
1260
ATTTAGGAAC TCGATACTGG TTTTTGTTAC TTTAGTTTAG CTTTTGTGAG GTTGAAACAA
1320
TTCAGATGTT TTTTTAACTT GTATATGTAA AGATCAATTT CTCGTGACAG TAAATAATAA
1380
TCCAATGTCT TCTGCCAAAT TAATATATGT ATTCGTATTT TTATATGAAA AAAAAAAAAA
1440
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA
1474

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1038 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GAATTCGGCA CGAGAGAGGG TTATATATCT TGATTCTGAC CTGATTGTCG TCGACGACAT
60
TGCCAAGCTC TGGGCCACGG ATTTGGAATC TCGTGTCCTC GGGGCACCAG AGTACTGCAA
120
GGCGAATTTC ACAAAGTATT TCACCGATAA TTTCTGGTGG GATCCCGCAT TATCCAAGAC
180
CTTTGAGGGA AAAAAACCCT GCTACTTCAA CACAGGCGTA ATGGTGATCG ATCTTGAAAA
240

ATGGCGGGCA GGGGAATTCA CAAGAAAGAT CGAAATCTGG ATGGACATAC AGAAGGAACG
300
CCGTATCTAT GAGCTCGGAT CATTACCGCC ATTTTTACTG GTATTTGCTG TTTTGGTTAA
360
GCAAGTCGAT CATCGTTGGA ATCAGCACGG TTTAGGCGGA GATAATTTGC AAGGCCTTTG
420
CCGAGATCTT CACCCTGGAC CTGTCAGTTT GTTGCATTGG AGTGGTAAGG GCAAACCTTG
480
GCTACGCCTG GAATGCCAAG CGGACTTGCC CTCTGGATAC TTTATGGGCT CCTTATGATC
540
TTTATCGATC AACGTATTAC CTAAATGGGT GAGAGAGCCT CTCTCCTCGG GGTGCTTTTT
600
ATCGAATTAA ACCTGATTTG ATAAAATGCC AAATAGAACT TTACGCCTAT GCATCTTTCA
660
GTTTTGAATT TCAATTCTGG TAACGAATAG AAGAAAACAA TAGCACAGCC ACAGGCAGGA
720
CAAATCCATC ATGAGGGACC AATCGTTTGA ATTTAGTATT AATAAGGTTG TTCCATATAA
780
CGCCTGTGAA GAATGATATT GTGGACTGAT CTATTTATAT TTGTACTGCC ATGCCATCCT
840
CAGCCAGCAG AGAGGCAAGC AATGCCGCTG CAAGTCATGT AGGGAAGGCG TTGTGAACTC
900
AATTTTCGGC GACTGTACAG GATGTAAATT TTTGGAACAT TAATATCATT ATGATAAGTT
960
CCTGAACCAA CAACTGTATA ATACCTTATA AATGTATCTG CAACTCCATT TTTGCATAAA
1020
AAAAAAAAAA AAAAAAAA
1038

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

CTAGGGGTCT TGGGGGGTTC CTGATGCCCA ATTGTTGCTG TGCTTGGCAT GAACCCAAAA
60
CATGCAAGAG ATCTGTAGTC AGTAGTCTTG TTGGATCTAT AGCTTTTAGA AAAGAGTCAC
120
GTCCTTTTAG GGTAACATCA TTCCAACCAT ATCCAGTTCC ACCACCGGCT ACACCTTCAA
180
CGGGAGGAGG AGCAAGATAT TCAGCATTGC TTTGGGCACC AGATGGATAG GCATTATTTT
240
CCATCGGAAT TCAGCCGAGC TCGCCCCCTC AGTCCAATCG TCGTGAAAAT CCCTCAAAAT
300
TGGGCAATTC TGGCTCGAAA TCGCCAAATT ATGGGCTACA ACAGGATTAA AATTGCACAG
360
AAATCTGCCA GT
372

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 545 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

AAAGAATTCG GCACGAGGGC AATCCGAGCC TAGCCAACCA ACTTGGCAGC AAGGAGCACA
60
GGGAGTTGGC GAGAGAAGCT GTTAGGAAAT CTTTGGTATT GTTGAAAAAT GGGAAGTCAG
120
CCAACAAGCC TTTGCTCCCT TTGGAGAAGA ATGCTTCCAA GGTTCTTGTT GCAGGAACCC
180
ATCCTGATAA TCTGGGTTAT CAGTGTGGTG GATGGACGAT GGAATGGCAA GGATTAAGTG
240
GAAACATAAC CGTAGGAACT ACAATTCTGG AAGCTATCAA ACTAGCTGTC AGCCCCTCTA
300
CTGAAGTGGT TTATGAGCAA AATCCAGATG CTAACTATGT CAAAGGACAA GGGTTTTCAT
360
ATGCCATTGT GGTTGTGGGT GAGGCACCAT ACGCAGAAAC GTTTGGAGAC AATCTTAATT
420
TGACCATTCC CCTAGGCGGA GGGGACACGA TTAAGACGGT CTGTGGCTCC TTGAAATGCC
480
TTGTAATCTT GATATCTGGA AGGCCACTTG TTATTGAACC TTATCTTCCA TTGGTGGATC
540
GTTTT
545

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 463 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

GCAGGTCGAC ACTAGTGGAT CCAAAGAATT CGGCACGAGA AAAAACAAAT GTTAGCTAGC
60
CTAGTGATGA GCTTTACGTA TACCTGGCCT TTTATACATG GATCTGAGTT TTTATGCAGG
120
TGTAGAGCCT TTTGTTACTC TGTATCACTG GGACTTGCCA CAAGCTCTGG AGGACGAATA
180
CGGTGGATTT CGTAGCAAAA AAGTTGTGGA TGACTTTGGC ATATTCTCAG AAGAATGCTT
240
TCGTGCTTTT GGAGACCGTG TGAAGTACTG GGTAACTGTT AACGAACCGT TGATCTTCTC
300
ATATTTTTCT TACGATGTGG GGCTTCACGC ACCGGGCCGC TGTTCGCCTG GATTTGGAAA
360
CTGCACTGCG GGAAATTCAG CGACAGAGCC TTATATTGTA GCCCATAACA TGCTTCTTGC
420
ACATAGTACC GCTGTTAAAA ATATATAGCA TAAATACCCA GGG
463

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 435 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

ACACTAGTGG ATCCAAAGAA TTCGGCACGA GGCTACCATC TTCCCTCATA ATATTGGGCT
60
TGGAGCTACC AGGGATCCTG ATCTGGCTAG AAGAATAGGG GCTGCTACGG CTTTGGAAGT
120
TCGAGCTACT GGCATTCAAT ACACATTTGC TCCATGTGTT GCTGTTTGCA GAGATCCTCG
180

ATGGGGCCGC TGCTATGAGA GCTACAGTGA GGATCCAAAA ATTTTCAAGG TTATGACTGA
240
GATTATCGTT GGCCTGCAAG GGAATCCTCC TGCTAATTCT ACAAAAGGGG GGCCTTTTAT
300
AGCTGGACAG TCAAATGTTG CAGCTTGTGC TAAGCATTTT GTGGGTTATG GTGGAACAAC
360
CAAAGGTATC GATGAGAATA ATACTGTTAT CAACTATCAA GGGTTATTTC AACATTCCAA
420
ATTACCCCCA ATTTT
435

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 451 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GAATTCGGCA CGAGCCTAGA ATTCTATGGT GAAAATTGTT GGGACAAGGC TGCCCAAGTT
60
TACAAAGGAA CAGTCCCAAA TGGTTAAAGG TTCAATAGAC TATCTAGGCG TTAACCAATA
120
CACTGCTTAT TACATGTATG ATCCTAAACA ACCTAAACAA AATGTAACAG ATTACCAGAC
180
TGGACTGGAA TACAGGCTTT GCATATGCTC GCAATGGAGT GCCTATTGGA CCAAGGGCGA
240
ACTCCAATTG GCTTTACATT GTGCCTTGGG GTCTATACAA GGCCGTCACA TACGTAAAAG
300
AACACTATGG AAATCCAACT ATGATTCTCT CTGAAAATGG AATGGACGAC CTGGAAACGT
360
GACACTTCCA GCAGGACTGC ATGATACCAT CAGGGGTAAC TACTATAAAA GCTATTTGCA
420
AAATTTGATT AATGCACGTG AATGACCGGG G
451

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 374 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CTGCTCTGCA AGCAGTACTA TGCACAGCAA GGCCTGCTTA ACTGAAAACA GAGCGCTGAG
60
CTTGAGGAAA CGCTCAAGCA TTGCTGAGGC CACCGTTTAT CTAAATAGCG CAACATAGGG
120
CTTCAGAAAA ATGGCAATGG CACAAGCATT CAGAGGCCGT GTCTTGCAAG CTGCCCGTTT
180
GCTCCGCCGC AACATTCTGC CGGAGGATAA AAGCTTTGGA TCCGCTGCTT CTCCTAGACG
240
AGCTCTTAGC CTGCTCTCAT CAAAAGCCTT CATCTCTTTC TCTGTTGAAC GGCATCGGCT
300
AGCTGCTACA AATTCAACAA TTGTGTTGCA ATCTCGAAAC TTTTCTGCAA AAGGTAAAAA
360
GACAGGACAA TCTG
374

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 457 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GAAGAATGGA AGAGATTAAT GGTGATAACG CAGTAAGGAG GAGCTGCTTT CCTCCAGGTT
60
TCATGTTTGG GATAGCAACT TCTGCTTATC AGTGTGAAGG AGCTGCCAAC GAAGGTGGAA
120
AAGGCCCAAG CATCTGGGAC TCATTTTCAC GAACACCAGG CAAAATTCTT GATGGAAGCA
180
ACGGTGATGT AGCAGTGGAT CAGTATCATC GTTATAAGGC AGATGTAAAA CTGATGAAAG
240
ATATGGGCGT GGCTACCTAC AGATTCTCGA TTTCATGGCC TCGTATATTT CCAAAGGGAA
300
AAGGAGAGAT CAATGAGGAA GGAGTAGCCT ATTACAATAA CCTCATCAAT GAACTCCTCC
360
AGAATGGAAT CCAAGCGTCT GTCAACTTTG TTTCACTGGG ATACTCCCCA GTCTCTGGAG
420
GATGAATATG GCGGATTTCT GAGGCCAACC ATTGTGA
457

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 346 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

GGTGTGATGG CAGGAATTCC AGTCCTAAGG CCATTTTGCA TCTGTTTGCT TTCAGTCTAC
60
ATGCTGCACA TTGTAGCTCC AGTAGCTTCA CCAAGGCTAG GTAGAAGCAG CTTCCCAAGG
120
GGTTTCAAAT TTGGTGCAGG GTCATCTGCT TATCAGGCGG AAGGAGCTGC TCATGAGGGT
180
GGCAAAGGCC CAAGCATTTG GGATACATTC TCCCACACTC CAGGTAAAAT CGCTGATGGG
240
AATATTGGGA TGTTGCAGTA GATCAATACC ACCGTTATAA GGAAGATGTG CAGCTTCTCA
300
AATACATGGG AATGGACGTC TATCGTTTCT CTATCTCCTG GTCACG
346

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 957 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

GAATTCGGCA CGAGAAAGCC CTAGAATTTT TTCAGCATGC TATCACAGCC CCAGCGACAA
60
CTTTAACTGC AATAACTGTG GAAGCGTACA AAAAGTTTGT CCTAGTTTCT CTCATTCAGA
120

CTGGTCAGGT TCCAGCATTT CCAAAATACA CACCTGCTGT TGTCCAAAGA AATTTGAAAT
180
CTTGCACTCA GCCCTACATT GATTTAGCAA ACAACTACAG TAGTGGGAAA ATTTCTGTAT
240
TGGAAGCTTG TGTCAACACG AACACAGAGA AGTTCAAGAA TGATAGTAAT TTGGGGTTAG
300
TCAAGCAAGT TTTGTCATCT CTTTATAAAC GGAATATTCA GAGATTGACA CAGACATATC
360
TGACCCTCTC TCTTCAAGAC ATAGCAAGTA CGGTACAGTT GGAGACTGCT AAGCAGGCTG
420
AACTCCATGT TCTGCAGATG ATTCAAGATG GTGAGATTTT TGCAACCATA AATCAGAAAG
480
ATGGGATGGT GAGCTTCAAT GAGGATCCTG AACAGTACAA AACATGTCAG ATGACTGAAT
540
ATATAGATAC TGCAATTCGG AGAATCATGG CACTATCAAA GAAGCTCACC ACAGTAGATG
600
AGCAGATTTC GTGTGATCAT TCCTACCTGA GTAAGGTGGG GAGAGAGCGT TCAAGATTTG
660
ACATAGATGA TTTTGATACT GTTCCCCAGA AGTTCACAAA TATGTAACAA ATGATGTAAA
720
TCATCTTCAA GACTCGCTTA TATTCATTAC TTTCTATGTG AATTGATAGT CTGTTAACAA
780
TAGTACTGTG GCTGAGTCCA GAAAGGATCT CTCGGTATTA TCACTTGACA TGCCATCAAA
840
AAAATCTCAA ATTTCTCGAT GTCTAGTCTT GATTTTGATT ATGAATGCGA CTTTTAGTTG
900
TGACATTTGA GCACCTCGAG TGAACTACAA AGTTGCATGT TAAAAAAAAA AAAAAAA
957

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 489 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GCAGGTCGAC ACTAGTGGAT CCAAAGAATT CGGCACGAGA TAAGACTAAT TTTCCAGACA
60
ATCCTCCATT CCCATTCAAT TACACTGGTA CTCCACCCAA TAATACACAG GCTGTGAATG
120
GGACTAGAGT AAAAGTCCTT CCCTTTAACA CAACTGTTCA ATTGATTCTT CAAGACACCA
180
GCATCTTCAG CACAGACAGC CACCCTGTCC ATCTCCATGG TTTCAATTTC TTTGTGGTGG
240
GCCAAGGTGT TGGAAACTAC AATGAATCAA CAGATGCACC AAATTTTAAC CTCATTGACC
300
CTGTCGAGAG AAACACTGTG GGAGTTCCCA AAGGAGGTTG GGCTGCTATA AGATTTCGTG
360
CAGACAATCC AGGGGTTTGG TTCATGCACT GTCATTTGGA GGTTCACACA TCGTGGGGAC
420
TGAAAATGGC GTGGGTAGTA AAGAACGGAA AAGGGCCCAT CGATTTTCCA CCCGGGTGGG
480
TACCAGTAA
489

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 471 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GAATTCGGCA CGAGAAAACC TTTTCAGACG AATGTTCTGA TGCTCGGCCC CGGCCAGACA
60
ACAGACATAC TTCTCACTGC CAATCAGGCT ACAGGTAGAT ACTACATGGC TGCTCGAGCA
120
TATTCCAACG GGCAAGGAGT TCCCTTCGAT AACACCACTA CCACTGCCAT TTTAGAATAC
180
GAGGGAAGCT CTAAGACTTC AACTCCAGTC ATGCCTAATC TTCCATTCTA TAACGACACC
240
AACAGTGCTA CTAGCTTCGC TAATGGTCTT AGAAGCTTGG GCTCACACGA CCACCCAGTC
300
TTCGTTCCTC AGAGTGTGGA GGAGAATCTG TTCTACACCA TCGGTTTGGG GTTGATCAAA
360
TGTCCGGGGC AGTCTTGTGG AGGTCCAACG GATCAAGATT TGCAGCAAGT ATGAATACAT
420
ATCATTTGTC CCGCAACCAC TTCTTCCAAT CCTTCAAGCT CAGCATTTTG G
471

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 338 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

GTTCGGCACT GAGAGATCCA TTTCTTTCAA TGTTGAGACA GTGAGTAGTA TTAGTTTGAT
60
ATCTCTTTCA GGAATATATC GTGCTTGCAG GATCTTTAGT TTCTGCAACA ATGTCGTTGC
120
AATCAGTGCG TCTATCTTCT GCTCTCCTTG TTTTGCTACT AGCATTTGTT GCTTACTTAG
180
TTGCTGTAAC AAACGCAGAT GTCCACAATT ATACCTTCAT TATTAGAAAG AGACAGTTAC
240
CAGGCTATGC AATAAGCGTA TAATCGCCAC CGTCAATGGC AGCTACCAGG CCCAACTATT
300
CATGTACGTG ATGGAGACGT TGTTAATTAT CAAAGCTT
338

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1229 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

AGAGAAATAA TTATATTTGT AAATTTAAGT CTACGTTTAT TAAAAAACTA CAACCCTAAA
60
TGCAGGAGAA AAAACAAGCA TGCTGTCTAC TGAAGCTTAC AAATCAAATC CCTGCGATAT
120
GTCTTTTCTC GTGCCGAATT CGGCACGAGA AGATCTTGGT TCGAGTCTCT CAGCTCTCTC
180
CAAAGGAATT TTGTGGGTCA TTTGCAGGTG AAGACACCAT GGTGAAGGCT TATCCCACCG
240
TAAGCGAGGA GTACAAGGCT GCCATTGACA AATGCAAGAG GAAGCTCCGA GCTCTCATTG
300

CAGAGAAGAA CTGTGCGCCG ATCATGGTTC GAATCGCATG GCACAGCGCT GGGACTTACG
360
ATGTCAAGAC CAAGACCGGA GGGCCCTTCG GGACGATGAG ATATGGGGCC GAGCTTGCCC
420
ACGGTGCTAA CAGTGGTCTG GACATCGCAG TTAGGCTCCT GGAGCCAATC AAGGAACAGT
480
TCCCCATAAT CACCTATGCT GACCTTTATC AGTTGGCTGG TGTGGTGGCT GTTGAAGTGA
540
CCGGGGGACC TGACATTCCG TTCCATCCTG GAAGAGAAGA CAAGCCTGAG CCTCCAGAAG
600
AAGGCCGCCT TCCTGATGCT ACAAAAGGAC CTGATCATCT GAGGGATGTT TTTGGTCACA
660
TGGGGTTGAA TGATAAGGAA ATTGTGGCCT TGTCTGGTGC CCACACCTTG GGGAGATGCC
720
ACAAGGAGAG ATCTGGTTTT GAAGGACCAT GGACCTCTAA CCCCCTTATC TTTGACAACT
780
CTTACTTCAC AGAGCTTGTG ACTGGAGAGA AGGAAGGCCT GCTTCAGTTG CCATCTGATA
840
AGGCACTGCT TGCTGATCCT AGTTTTGCAG TTTATGTTCA GAAGTATGCA CAGGACGAAG
900
ACGCTTTCTT TGCTGACTAT GCGGAAGCTC ACCTGAAGCT TTCTGAACTT GGGTTTGCTG
960
ATGCGTAGAT TCATACCTTC TGCAGAGACA ATTCCTTGCT AGATAGCTTC GTTTTGTATT
1020
TCATCTAATC TTTTCGATTA TATAGTCACA TAGAAGTTGG TGTTATGCGC CATAGTGATA
1080
CTTGAACCTA CATGTTTTTG AAAAGTATCG ATGTTCTTTA AAATGAACAT TGAATACAAC
1140
ATTTTGGAAT CTGGTTGTGT TCTATCAAGC GCATATTTTA ATCGAATGCT TCGTTCCTGT
1200
TAAAAAAAAA AATAAAATAA AAAAAAAAA
1229

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1410 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

GAAGATGGGG CTGTGGGTGG TGCTGGCTTT GGCGCTCAGT GCGCACTATT GCAGTCTCAG
60
GCTTACAATG TGGTAAGTTC AAGCAATGCT ACTGGGAGTT ACAGTGAGAA TGGATTGGTG
120
ATGAATTACT ATGGGGACTC TTGCCCTCAG GCTGAAGAGA TCATTGCTGA ACAAGTACGC
180
CTGTTGTACA AAAGACACAA GAACACTGCA TTCTCATGGC TTAGAAATAT TTTCCATGAC
240
TGTGCTGTGG AGTCATGTGA TGCATCGCTT CTGTTGGACT CAACAAGGAA CAGCATATCA
300
GAAAAGGACA CTGACAGGAG CTTCGGCCTC CGCAACTTTA GGTATTTGGA TACCATCAAG
360
GAAGCCGTGG AGAGGGAGTG CCCCGGGGTC GTTTCCTGTG CAGATATACT CGTTCTCTCT
420
GCCAGAGATG GCGTTGTATC GTTGGGAGGA CCATACATTC CCCTGAAGAC GGGAAGAAGA
480
GATGGACGGA AGAGCAGAGC AGATGTGGTG GAGAATTACC TGCCCGATCA CAATGAGAGC
540
ATCTCCACTG TTCTGTCTCG CTTCAAAGCC ATGGGAATCG ACACCCGTGG GGTTGTTGCA
600
CTGCTGGGGG CTCACAGCGT GGGGAGGACT CACTGCGTGA AGCTGGTGCA CAGGCTGTAC
660

CCGGAAGTAG ATCCGACACT GGACCCTGGG CACGTGGAGC ACATGAAGCA CAAGTGCCCC
720
GACGCGATCC CCAACCCGAA GGCAGTGCAG TATGTGCGGA ACGACCGGGG AACGCCTATG
780
AAGCTGGACA ACAACTACTA CGTGAACCTG ATGAACAAGA AGGGGCTCCT AATAGTGGAC
840
CAGCAACTGT ATGCAGATTC GAGGACCAGG CCGTATGTGA AGAAGATGGC AAAAAGCCAG
900
GAATACTTCT TCAAATACTT CTCCCGGGGC CTCACCATCC TCTCTGAGAA CAATCCTCTC
960
ACCGGCGCTC GAGGAGAAAT CCGTCGGCAG TGCTCGCTCA AAAACAAATT CCACACAAAA
1020
AGCAAGCGTT GAGCGATAGC TCAATGCCGC AGTGGTGGGA GTGATAGCGT GATGCCACAG
1080
TGGTGGGCAT TTCATATATA AATTGCAGTT TGCGTTTTTA TTAGATAATC ATAATGGTGT
1140
GGTGTGACTA TGCCCTGCGA ATCACATCGA TGAACCACAA CCGAACCGTG GAACAGTAGG
1200
CTTATTCCCT TATGTAAGCA GAACCTTTTA TTATAAGCAA AAAAGACAAT CCTGTCTCTT
1260
ATTCTAGTAT AATTTTGTCA TCAGTTAAAG TTGCTCATCT GATAATAACT GGAAACGGTA
1320
AAATATGACA ACTACGTATC TTCTTTGGTC ATCTGATAAT AACCGGAAAC GATAAAATAT
1380
GACAACTACA TATATTCTTT AAAAAAAAAA
1410

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 687 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GTAGTTTCGT TTTACAACAA TCTCAGGTTT TGAATCTCAG AATAGTTGCG AAAGGAAGCG
60
ATGACGAAGT ACGTGATCGT TAGCTCCATT GTGTGTTTCT TTGTATTTGT TTCTGCGTGC
120
ATAATTTCTG TCAATGGATT AGTTGTCCAT GAAGATGATC TGTCAAAGCC TGTGCATGGG
180
CTTTCGTGGA CATTTTATAA GGACAGTTGC CCCGACTTGG AGGCCATAGT GAAATCGGTA
240
CTTGAGCCGG CGTTGGACGA AGATATCACT CAGGCCGCAG GCTTGCTGAG ACTTCATTTC
300
CATGACTGTT TTGTGCAGGG TTGCGATGGG TCCGTGTTGC TGACAGGAAC TAAAAGAAAC
360
CCCAGTGAGC AACAGGCTCA GCCAAACTTA ACACTAAGAG CCCGGGCCTT GCAGCTGATC
420
GACGAAATTA AAACCGCTGT AGAAGCTAGC TGCAGTGGGG TTGTAACTTG TGCAGACATT
480
CTGGCTTTGG CTGCTCGTGA CTCCGTCCGC TCAGGAGGCC CAAAATTTCC AGTACCACTT
540
GGCCGCAGAG ATAGCCTAAA GTTTGCCAGT CAATCCGTAG TTGTCGCCAA TATACCAACT
600
CCAACTTTAA ATTTGACACA GCTGATGAAC ATTTTTGGCT TCAAAGGATT CAGTTTGGCC
660
GAAATGGTTG CTCTTCAGGT GGCACAC
687

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 688 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

GTAGTTTCGT TTTACAACAA TCTACAGGTT TTGAATCTCA GAATAGTTGC GAAAGGAAGC
60
GATGACGAAG TACGTGATCG TTAGCTCCAT TGTATGTTTC TTTGTATTTG TTTCTCCGTG
120
CATAATTTCT GTCAATGGAT TAGTTGTCCA TGAAGATGAT CTGTCAAAGC TTGTGCATGG
180
GCTTTCGTGG ACATTTTATA AGGACAGTTG CCCCGACTTG GAGGCCATAG TGAAATCGGT
240
ACTTGAGCCG GCGTTGGACG AAGATATCAC TCAGGCCGCA GGTTGCTGAG ACTTCATTTC
300
CATGACTGTT TTGTGCAGGG TTGCGATGGG TCCGTGTTGC TGACAGGAAC TAAAAGAAAC
360
CCCCGAGTGA GCAACAGGCT CAGCCAAACT TAACACTAAG AGCCCGGGCC TTGCAGCTGA
420
TCGACGAAAT TAAAACCGCT GTAGAAGCTA GCTGCAGTGG GGTTGTAACT TGTGCAGACA
480
TTGTGGCTTT GGCTGCTCGT GACTCCGTCG CTCAGGAGGC GCAAAATTTC GAGTACCAGT
540
TGGCCGCAGA GATAGCCTAA AGTTTGCCAG TCAATCCGTA GTTGTCGCCA ATATACCAAC
600
TCCAACTTTA AATTTGACAC AGCTGATGAA CATTTTTGGC TCCAAAGGAT TCAGTTTGGC
660
CGAAATGGTT GCTCTTCAGG TGGCACAC
688

Claims:

1. An isolated DNA sequence comprising a nucleotide sequence selected from the group consisting of

(a) sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88;

(b) complements of the sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88;

(c) reverse complements of the sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88;

(d) reverse sequences of the sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and

(e) sequences having at least about a 99% probability of being the same as a sequence of (a) - (d) as measured by computer algorithm FASTA.

2. A DNA construct comprising a DNA sequence according to claim 1.

3. A transgenic cell comprising a DNA construct according to claim 2.

4. A DNA construct comprising, in the 5'-3' direction:

(a) a gene promoter sequence,

(b) an open reading frame coding for at least a functional portion of an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

5. The DNA construct of claim 4 wherein the open reading frame is in a sense orientation.

6. The DNA construct of claim 4 wherein the open reading frame is in an antisense orientation.

7. The DNA construct of claim 4, wherein the gene promoter sequence and gene termination sequences are functional in a plant host.

8. The DNA construct of claim 4, wherein the gene promoter sequence provides for transcription in xylem.

9. The DNA construct of claim 4 further comprising a marker for identification of transformed cells.

10. A DNA construct comprising, in the 5'-3' direction:

(a) a gene promoter sequence,

(b) a non-coding region of a gene coding for an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

11. The DNA construct of claim 10 wherein the non-coding region is in a sense orientation.

12. The DNA construct of claim 10 wherein the non-coding region is in an antisense orientation.

13. The DNA construct of claim 10, wherein the gene promoter sequence and gene termination sequences are functional in a plant host.

14. The DNA construct of claim 10, wherein the gene promoter sequence provides for transcription in xylem.

15. A transgenic plant cell comprising a DNA construct, the DNA construct comprising, in the 5'-3' direction:

(a) a gene promoter sequence;

(b) an open reading frame coding for at least a functional portion of an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

16. The transgenic plant cell of claim 15 wherein the open reading frame is in a sense orientation.

17. The transgenic plant cell of claim 15 wherein the open reading frame is in an antisense orientation.

18. The transgenic plant cell of claim 15 wherein the DNA construct further comprises a marker for identification of transformed cells.

19. A plant comprising a transgenic plant cell according to claim 15, or fruit or seeds thereof.

20. The plant of claim 19 wherein the plant is a woody plant.

21. The plant of claim 20 wherein the plant is selected from the group consisting of eucalyptus and pine species.

22. A transgenic plant cell comprising a DNA construct, the DNA construct comprising, in the 5'-3' direction:

(a) a gene promoter sequence;

(b) a non-coding region of a gene coding for an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

23. The transgenic plant cell of claim 22 wherein the non-coding region is in a sense orientation.

24. The transgenic plant cell of claim 22 wherein the non-coding region is in an antisense orientation.

25. A plant comprising a transgenic plant cell according to claim 22, or fruit or seeds thereof.

26. The plant of claim 25 wherein the plant is a woody plant.

27. The plant of claim 26, wherein the plant is selected from the group consisting of eucalyptus and pine species.

28. A method for modulating the lignin content of a plant comprising stably incorporating into the genome of the plant a DNA construct comprising, in the 5`-3` direction:

(a) a gene promoter sequence;

(b) an open reading frame coding for at least a functional portion of an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

29. The method of claim 28 wherein the plant is selected from the group consisting of eucalyptus and pine species.

30. The method of claim 28 wherein the open reading frame is in a sense orientation.

31. The method of claim 28 wherein the open reading frame is in an antisense orientation.

32. A method for modulating the lignin content of a plant comprising stably incorporating into the genome of the plant a DNA construct comprising, in the 5'-3' direction:

(a) a gene promoter sequence;

(b) a non-coding region of a gene coding for an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

33. The method of claim 32 wherein the non-coding region is in a sense orientation.

34. The method of claim 32 wherein the non-coding region is in an antisense orientation.

35. The method of claim 32 wherein the plant is a woody plant.

36. The method of claim 35, wherein the plant is selected from the group consisting of eucalyptus and pine species.

37. A method for producing a plant having altered lignin structure comprising:

(a) transforming a plant cell with a DNA construct comprising, in the 5'-3' direction, a gene promoter sequence, an open reading frame coding for at least a functional portion of an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA, and a gene termination sequence to provide a transgenic cell;

(b) cultivating the transgenic cell under conditions conducive to regeneration and mature plant growth.

38. The method of claim 37 wherein the open reading frame is in a sense orientation.

39. The method of claim 37 wherein the open reading frame is in an antisense orientation.

40. The method of claim 37 wherein the plant is a woody plant.

41. The method of claim 40 wherein the plant is selected from the group consisting of eucalyptus and pine species.

42. A method for producing a plant having altered lignin structure comprising:

(a) transforming a plant cell with a DNA construct comprising, in the 5'-3' direction, a gene promoter sequence, a non-coding region of a gene coding for an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA, and a gene termination sequence to provide a transgenic cell:

(b) cultivating the transgenic cell under conditions conducive to regeneration and mature plant growth.

43. The method of claim 42 wherein the non-coding region is in a sense orientation.

44. The method of claim 42 wherein the non-coding region is in an antisense orientation.

45. The method of claim 42 wherein the plant is a woody plant.

46. The method of claim 45 wherein the plant is selected from the group consisting of eucalyptus and pine species.

47. A method of modifying the activity of an enzyme in a plant comprising stably incorporating into the genome of the plant a DNA construct including

(a) a gene promoter sequence;

(b) an open reading frame coding for at least a functional portion of an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

48. The method of claim 47 wherein the open reading frame is in a sense orientation.

49. The method of claim 47 wherein the open reading frame is in an antisense orientation.

50. A method of modifying the activity of an enzyme in a plant comprising stably incorporating into the genome of the plant a DNA construct including

(a) a gene promoter sequence;

(b) a non-coding region of a gene coding for an enzyme encoded by a nucleotide sequence selected from the group consisting of sequences recited in SEQ ID NO: 3, 13, 16-70, 72-88 and sequences having at least about a 99% probability of being the same as a sequence of SEQ ID NO: 3, 13, 16-70, 72-88 as measured by computer algorithm FASTA; and

(c) a gene termination sequence.

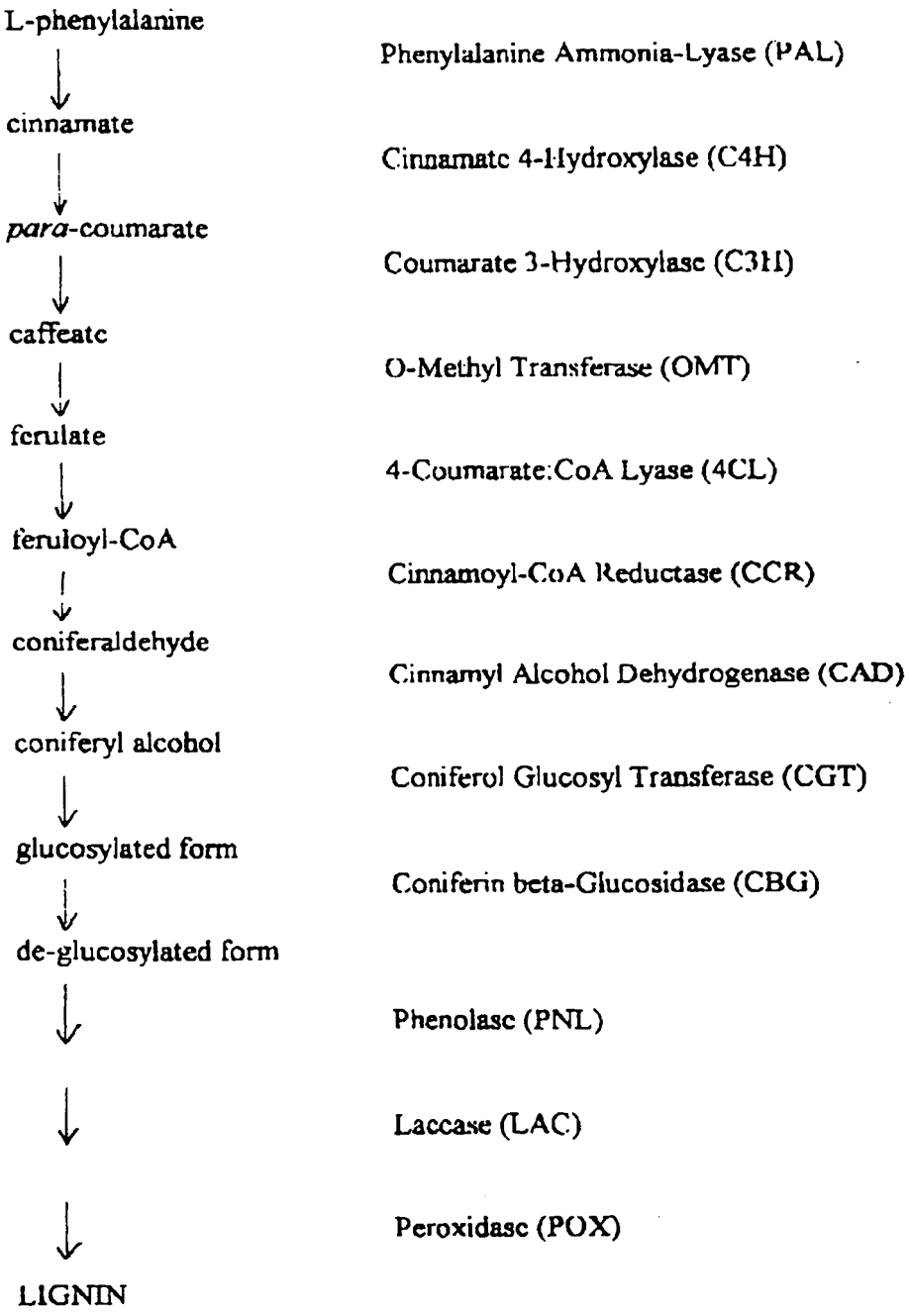
51. The method of claim 50 wherein the non-coding region is in a sense orientation.

52. The method of claim 50 wherein the non-coding region is in an antisense orientation.

53. The method of claim 50 wherein the plant is a woody plant.

54. The method of claim 53 wherein the plant is selected from the group consisting of eucalyptus and pine species.

FIG. 1


L-phenylalanine
Phenylalanine Ammonia-Lyase (PAL)
cinnamate
Cinnamate 4-Hydroxylase (C4H)
para-coumarate
Coumarate 3-Hydroxylase (C3H)
caffeate
O-Methyl Transferase (OMT)
ferulate
4-Coumarate:CoA Lyase (4CL)
feruloyl-CoA
Cinnamoyl-CoA Reductase (CCR)
coniferaldehyde
Cinnamyl Alcohol Dehydrogenase (CAD)
coniferyl alcohol
Coniferol Glucosyl Transferase (CGT)
glucosylated form
Coniferin beta-Glucosidase (CBG)
de-glucosylated form
Phenolase (PNL)
Laccase (LAC)
Peroxidase (POX)
LIGNIN

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/53, 15/54, 15/52, 15/60, 15/82, A01H 5/00

A3

(11) International Publication Number: **WO 98/11205**

(43) International Publication Date: 19 March 1998 (19.03.98)

(21) International Application Number: PCT/NZ97/00112

(22) International Filing Date: 10 September 1997 (10.09.97)

(30) Priority Data:
08/713,000 11 September 1996 (11.09.96) US

(71) Applicants: GENESIS RESEARCH & DEVELOPMENT CORPORATION LIMITED [NZ/NZ]; 1 Fox Street, Parnell, Auckland (NZ). FLETCHER CHALLENGE FORESTS LIMITED [NZ/NZ]; 585 Great South Road, Penrose, Auckland (NZ).

(72) Inventors: BLOKSBERG, Leonard, Nathan; 5A Korau Road, Greenlane, Auckland (NZ). GRIERSON, Alistair, Wallace; 1/24 Medina Place, Bucklands Beach, Auckland (NZ). HAVUKKALA, Ilkka, Jaakko; 3/121 Atkin Avenue, Mission Bay, Auckland (NZ).

(74) Agents: BENNETT, Michael, Roy et al.; Russell McVeagh West-Walker, The Todd Building, Level 5, 171-177 Lambton Quay, Wellington 6001 (NZ).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
20 August 1998 (20.08.98)

(54) Title: MATERIALS AND METHODS FOR THE MODIFICATION OF PLANT LIGNIN CONTENT

(57) Abstract

Novel isolated DNA sequences associated with the lignin biosynthetic pathway are provided, together with DNA constructs including such sequences. Methods for the modulation of lignin content in plants are also disclosed, the methods comprising incorporating one or more of the inventive DNA sequences or a sequence complementary to an inventive DNA sequence into the genome of a plant.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Interna[illegible]plication No
PCT/NZ 97/00112

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/53 C12N15/54 C12N15/52 C12N15/60 C12N15/82
A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BOUDET A M ET AL: "TANSLEY REVIEW NO. 80 BIOCHEMISTRY AND MOLECULAR BIOLOGY OF LIGNIFICATION" NEW PHYTOLOGIST, vol. 129, no. 2, 1 January 1995, pages 203-236, XP002006037 see page 216 - page 229 ---	1-54
X	BOUDET A M ET AL: "LIGNIN GENETIC ENGINEERING" MOLECULAR BREEDING, vol. 2, 1996, pages 25-39, XP002025844 see the whole document --- -/--	1-54

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 May 1998	07.07.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Maddox, A

4

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
O,X	DIXON, R.A., ET AL: "Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses - a review." GENE. PAPERS PRESENTED AT THE CHULABHORN RESEARCH INSTITUTE CONFERENCE HELD BANGKOK, THAILAND, 7-10 AUGUST 1995., vol. 179, November 1996, pages 61-71, XP002054131 see page 65, left-hand column, last paragraph	1-7
X	HOTZE, M., ET AL.: "Cinnamate 4-hydroxylase from Catharanthus roseus, and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in Escherichia coli" FEBS LETTERS, vol. 374, 1995, pages 345-350, XP002054132	1-5,9-11
Y	see the whole document & HOTZE, M., ET AL.: "C.roseus mRNA for cinnamate 4-hydroxylase (CYP73)" EMBL SEQUENCE DATABASE, REL.39, 15-APR-1994, ACCESSION NO. Z32563, XP002054206	6-8, 12-54
Y	EP 0 716 147 A (JUJO PAPER CO LTD) 12 June 1996 see page 8, line 17	6-8, 12-54
X	MIZUTANI, M., ET AL.: "Molecular clonng and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate-4-hydroxylase activity" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 190, no. 3, 1993, pages 875-880, XP002054134 see the whole document	1-3
X	KAWAI, S., ET AL.: "Populus kitakamiensis cyp73a gene for cinnamic acid 4-hydroxylase complete cds." EMBL SEQUENCE DATABASE, REL.46 30-DEC-1995, ACCESSION NO. D82812, XP002054135 see the whole document	1-3

-/--

4

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 632 128 A (KYOWA HAKKO KOGYO KK) 4 January 1995 see sequence ID no.63 see example 13	1-4
X	WO 95 07993 A (ZENECA LTD ;SMART CATHERINE MARGARET (GB); THOMAS HOWARD (GB); HOS) 23 March 1995 see sequence ID13	1-3
X A	WO 93 05160 A (ICI PLC) 18 March 1993 see figure 1	1-3 10-54
X	EP 0 516 958 A (BAYER AG) 9 December 1992 see the whole document	1-3
X	DOORSSELAERE VAN J ET AL: "A NOVEL LIGNIN IN POPLAR TREES WITH A REDUCED CAFFEIC ACID/5-HYDROXYFERULIC ACID O-METHYLTRANSFERASE ACTIVITY" PLANT JOURNAL, vol. 8, no. 6, 1995, pages 855-864, XP002036470 see the whole document	1-4,6,7, 15-18, 37,42, 47,50
X	ATANASSOVA , R., ET AL: "Altered lignin composition in transgenic tobacco expressing O-methyltransferase sequences in sense and antisense oritentation" THE PLANT JOURNAL, vol. 8, no. 4, 1995, pages 465-477, XP002066043 see the whole document	1-7, 15-18, 28,30, 31, 37-39, 47-49
X	WO 94 23044 A (SAMUEL ROBERTS NOBLE FOUNDATIO) 13 October 1994 see the whole document	1-7,28, 32
X	NI, W., ET AL.: "Reduced lignin in transgenic plants containing a caffeic acid O-methyltransferase antisense gene" TRANSGENIC RESEARCH, vol. 3, 1994, pages 120-126, XP002054140 see the whole document	1-4,6,7, 28,32
X	HALPIN, C., ET AL.: "Manipulation of lignin quality by down-regulation of cinnamyl alcohol dehydrogenase" THE PLANT JOURNAL, vol. 6, 1994, pages 339-350, XP002054141 see the whole document	1-4,6,7, 37,42

-/--

4

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
O,X	CHABBERT ET AL: "Manipulation of lignin quality in transgenic poplar" BIOTECHNOL. PULP PAP. IND., PROC. INT. CONF., 6TH 1995, 1995, pages 319-322, XP002065202 see the whole document	1-4,6,7, 37,40, 42,45
X	BAUCHER, M., ET AL.: "Higher extractability of lignin in poplar (Populus tremula x P.alba) by reducing cinnamyl alcohol dehydrogenase activity" SOMATIC CELL GENETICS AND MOLECULAR GENETICS OF TREES. ISBN: 0-7923-4179-1., 1996, pages 153-158, XP002065203 see the whole document	1-4,6,7, 37,40, 42,45
X	WO 93 05159 A (ICI PLC) 18 March 1993	1-8,37, 42
A	see the whole document	10-54
X	WO 95 27790 A (CENTRE NAT RECH SCIENT ;BOUDET ALAIN (FR); PETTENATI JACQUELINE (F) 19 October 1995	1-3
A	see the whole document	10-54
X	BOUDET, A.M., ET AL.: "La lignification domestiquée" BIOFUTUR, vol. 158, July 1996, pages 27-31, XP002065204 see page 30, column 2 - column 3	1-4,6,7, 28,32, 37,42
O,X	BOUDET A M: "GENES INVOLVED IN MONOLIGNOL BIOSYNTHESIS AND THEIR MANIPULATION FOR TAILORING NEW LIGNINS" AMERICAN CHEMICAL SOCIETY. ABSTRACTS OF PAPER. AT THE NATIONAL MEETING, no. 1, 1996, XP000618488 see the whole document	1-4,6,7, 28,32, 37,42
X	WO 94 08036 A (SALK INST FOR BIOLOGICAL STUDI ;UNIV OLDENBERG (DE); CLAUSEN MONIK) 14 April 1994 see the whole document	1-7,28, 32

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ELKIND, Y., ET AL.: "Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 87, November 1990, WASHINGTON US, pages 9057-9061, XP002065205 see the whole document	1-7,28, 30,32,33
X	BATE, N.J., ET AL.: "Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determinig step in natural product biosynthesis" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 91, August 1994, WASHINGTON US, pages 7608-7612, XP002065206 see the whole document	1-7,28, 30,32,33
X	KAJITA S ET AL: "Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate:Coenzyme A ligase." PLANT AND CELL PHYSIOLOGY 37 (7). 1996. 957-965. , XP002065207 see the whole document	1-7,28, 30,37,42
X	EP 0 513 884 A (MOGEN INT) 19 November 1992 see page 4, line 31 - line 34	1-4,6, 15,17
X	WO 94 21794 A (ZENECA LTD ;ABU BAKAR UMI KALSOM (MY); BARTON SARAH LOUISE (GB); G) 29 September 1994 see page 47	1-7
X	WO 90 08828 A (PALADIN HYBRIDS INC) 9 August 1990 see figure 3C	1-4,6
O,X	ERIKSSON, K.-E. L. ET AL: "Laccase as a target for decreasing lignin content in transgenic trees through antisense genetic engineering" BIOTECHNOL. PULP PAP. IND., PROC. INT. CONF., 6TH (1996), MEETING DATE 1995, 310-314. EDITOR(S): SREBOTNIK, EWALD;MESSNER, KURT. PUBLISHER: FACULTAS-UNIVERSITAETSVERLAG, VIENNA, AUSTRIA. , XP002065208 see the whole document	1-5,37, 42

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LAGRIMINI L M: "WOUND-INDUCED DEPOSITION OF POLYPHENOLS IN TRANSGENIC PLANTS OVEREXPRESSING PEROXIDASE" PLANT PHYSIOLOGY, vol. 96, 1991, pages 577-583, XP002025778 see the whole document	1-4,6, 28,32
X	LIU, T.-T.Y., ET AL.: "Lignin content and composition in tobacco plants with over and under expressed peroxidase" SUPPLEMENT TO PLANT PHYSIOLOGY, vol. 102, no. 1, May 1993, page 103 XP002065209 see abstract 579	1-6,28, 32
X	MCINTYRE, C.L., ET AL.: "Strategies for the suppression of peroxidase gene expression in tobacco. II. In vivo suppression of peroxidase activity in transgenic tobacco using ribozyme and antisense constructs" TRANSGENIC RESEARCH, vol. 5, July 1996, pages 263-270, XP002065262 see the whole document	47,50
X	SIKORSKI, R.S., ET AL.: "Yeast centromere vector pRS415 with LEU2 marker, complete sequence" EMB SEQUENCE ACCESSION NO. U03449, 8 January 1984, XP002065210 see the whole document	1
X	YU, L.X., ET AL.: "Lycopersicon chilense unknown protein (LC15) mRNA, complete cds." EMBL ACCESSION NO. U19099, 3 October 1995, XP002065211 see the whole document	1
X	GRIMA-PETTENATI, J., ET AL.: "E.gunnii OMT mRNA for O-methyltransferase" EMBL ACCESSION NO.X74814, 31 December 1993, XP002065212 see the whole document -& POEYDOMENGE, O., ET AL.: "A cDNA encoding S-adenosyl-L-Methionine:caffeic acid 3-O-methyltransferase from Eucalyptus" PLANT PHYSIOLOGY, vol. 105, 1994, pages 749-750, XP002054139	1

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAYNAL, M., ET AL.: "A. thaliana transcribed sequence; clone PAP790; 5'end; Similar to cinnamyl alcohol dehydrogenase; Stylosanthes hmilis" EMBL ACCESSION NO. Z46703, 18 November 1994, XP002065213 see the whole document ---	1
X	GOFFNER, D., ET AL.: "E.gunnii mRNA for cinnamyl alcohol dehydrogenase" EMBL ACCESSION NO. X88797, 31 December 1995, XP002065214 see the whole document ---	1
X	NEWMAN, T., ET AL.: "10030 Arabidopsis thaliana cDNA clone 143C13T7" EMBL ACCESSION NO. T46767, 11 February 1995, XP002065215 see the whole document ---	1
X	ZHANG, X.H., ET AL.: "Pinus taedae phenylalanine ammonia-lyase (lpPAL) gene complete cds." EMBL ACCESSION NO. U39792, 1 January 1996, XP002065216 see the whole document ---	1
X	VOO, K.S., ET AL.: "Pinus taeda PT4CL2 4-coumarate-CoA ligase enzyme, mRNA, complete cds." EMBL ACCESSION NO. U12013, 27 July 1994, XP002065217 see the whole document ---	1
X	ZHANG, X.H., ET AL.: "Pinus taeda xylem 4-coumarate:CoA ligase (lp4CL-1) gene, complete cds." EMBL ACCESSION NO. U39405, 1 January 1996, XP002065218 see the whole document ---	1
X	DAVIES, K.M., ET AL.: "Malus sp. mRNA for anthocyanin hydroxylase" EMBL ACCESSION NO. X71360, 27 April 1993, XP002065219 see the whole document ---	1
X	HRMOVA, M., ET AL.: "Hordeum vulgare beta-d-glucan exohydrolase, isoenzyme exoII, mRNA, complete cds." EMBL ACCESSION NO. U46003, 29 February 1996, XP002065220 see the whole document ---	1

-/--

4

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WILLEKENS, H.D.: "N. plumbaginifolia mRNA for catalase (cat3 gene)" EMBL ACCESSION NO. Z36977, 7 September 1994, XP002065221 see the whole document	1
X	RITTER, D., ET AL.: "Gossypium hirsutum peroxidase mRNA, complete cds." EMBL ACCESSION NO. L08199, 24 December 1992, XP002065222 see the whole document	1
X	MEYER, K., ET AL.: "Arabidopsis thaliana ferulate-5-hydroxylase (FAH1) mRNA, complete cds." EMBL SEQUENCE ACCESSION NO. U38416, 13 August 1996, XP002065223 see the whole document -& MEYER, K., ET AL.: "Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 93, July 1996, WASHINGTON US, pages 6869-6874, XP002036466	1
P,X	SEWALT, V.J.H., ET AL.: "Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase" PLANT PHYSIOLOGY, vol. 115, September 1997, pages 41-50, XP002054136 see the whole document	1-13, 15-19, 22-25, 28, 30-34, 47-52
P,X	WO 97 23599 A (DU PONT ;PURDUE RESEARCH FOUNDATION (US); CHAPPLE CLINT (US)) 3 July 1997 see the whole document	1-5,37, 42
P,X	RECH, P., ET AL.: "E.gunnii mRNA for caffeoyl-CoA O-methyltransferase" EMBL SEQUENCE ACCESSION NO. Y12228, 8 April 1997, XP002065224 see the whole document	1

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE WPI Section Ch, Week 9737 Derwent Publications Ltd., London, GB; Class C06, AN 97-397027 XP002065227 & JP 09 173 069 A (MITSUBISHI PAPER MILLS LTD) see abstract	1-7,28, 32
A	WAGNER, A., ET AL.: "Pinus radiata cinnamyl alcohol dehydrogenase (CAD) mRNA, complete cds." EMBL SEQUENCE DATABASE, REL. 48, 28-JUL-1996, ACCESSSION NO. U62394, XP002054137 see the whole document	1
A	MASON, M.E., ET AL.: "Pinus elliottii PEC18 mRNA,partial sequence" EMBL SEQUENCE DATABASE, REL.47 31-MAY-1996, ACCESSION NO. U55006, XP002054138 see the whole document	1
A	BACHEM, C.W.B., ET AL.: "Antisense expression of polyphenol oxidase genes inhbits enzymatic browning in potato tubers" BIOTECHNOLOGY, vol. 12, November 1994, pages 1101-1105, XP002065225 see the whole document	1-3
A	WO 93 15599 A (CORNELL RES FOUNDATION INC) 19 August 1993 see the whole document	1-3
A	UDAGAMA-RANDENIYA,P.V., ET AL.: "Coniferyl alcohol oxidase: A catechol oxidase?" TREES, vol. 10, no. 2, 1995, pages 102-107, XP002065980 see the whole document	1
A	DHARMAWARDHANA, D.P., ET AL.: "A beta-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin" PLANT PHYSIOLOGY, vol. 107, 1995, pages 331-339, XP002065226 see the whole document	1

-/--

4

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE DISSABS AN 97:45741 DISSABS Order Number: AARNN14739, DHARMAWARDHANA, D.P., ET AL.: "A biochemical and molecular study of lignin biosynthesis (Pinus contorta, glucosidase, coniferin, xylem)" XP002065982 see abstract & PH.D THESIS, UNIVERSITY OF BRITISH COLUMBIA AVAILABLE FROM DISSERTATION ABSTRACTS INTERNATIONAL, ORDER NUMBER AARNN14739, ISBN: 0-612-14739-8, 1996, ---	1
A	BAO, W., ET AL.: "A laccase associated with lignification in loblolly pine xylem." SCIENCE, vol. 260, 30 April 1993, pages 672-674, XP002065981 see the whole document ---	1
A	CHEMICAL ABSTRACTS, vol. 125, no. 13, 23 September 1996 Columbus, Ohio, US; abstract no. 163462, SHIOKAWA, T., ET AL.: "Expression analysis of a cinnamic acid 4-hydroxylase gene from a hybrid aspen, Populus kitakamiensis" XP002054142 see abstract & KAMI PARUPU KENKYU HAPPYOKAI KOEN YOSHISHU, vol. 63rd, 1996, pages 160-163, ---	10
A	WO 93 24638 A (ZENECA LTD ;WALTER MICHAEL HERBERT (DE)) 9 December 1993 see the whole document ---	8,14
A	WO 96 20595 A (PROGUARD INC ;EMERSON RALPH W (US); CRANDALL BRADFORD G JR (US)) 11 July 1996 see page 20 - page 22; example 9 ---	6,12,17, 24,31, 34,39, 44,49,52
E	WO 97 45549 A (CENTRE NAT RECH SCIENT ;FAYE LOIC (FR); GOMORD VERONIQUE MARTINE () 4 December 1997 see sequence IDs 3 and 5 -----	1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/NZ 97/00112

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0716147 A	12-06-1996	JP 9154580 A	17-06-1997
		AU 3855795 A	06-06-1996
		BG 101524 A	30-01-1998
		CA 2162449 A	10-05-1996
		CN 1137565 A	11-12-1996
		CZ 9701388 A	18-02-1998
		FI 971961 A	07-07-1997
		HU 77074 A	02-03-1998
		WO 9615252 A	23-05-1996
		NO 972108 A	07-07-1997
		PL 320201 A	15-09-1997
EP 0632128 A	04-01-1995	AU 680401 B	31-07-1997
		CA 2130800 A	03-09-1993
		WO 9318155 A	16-09-1993
WO 9507993 A	23-03-1995	AU 7619494 A	03-04-1995
		CA 2172842 A	23-03-1995
		EP 0719341 A	03-07-1996
WO 9305160 A	18-03-1993	AU 663726 B	19-10-1995
		AU 2516792 A	05-04-1993
		BR 9206481 A	31-10-1995
		EP 0603250 A	29-06-1994
		JP 6510429 T	24-11-1994
EP 0516958 A	09-12-1992	DE 4117747 A	03-12-1992
		CA 2067317 A	01-12-1992
		JP 5199886 A	10-08-1993
		US 5728570 A	17-03-1998
WO 9423044 A	13-10-1994	AU 6621194 A	24-10-1994
WO 9305159 A	18-03-1993	AU 669106 B	30-05-1996
		AU 1658192 A	05-04-1993
		BR 9205934 A	05-07-1994
		CA 2109222 A	27-10-1992
		EP 0584117 A	02-03-1994
		JP 6509465 T	27-10-1994
		US 5451514 A	19-09-1995

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/NZ 97/00112

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9527790 A	19-10-1995	FR 2718460 A	13-10-1995
		AU 2347295 A	30-10-1995
		BR 9507291 A	23-09-1997
		CA 2185334 A	19-10-1995
		EP 0755449 A	29-01-1997
		JP 9511647 T	25-11-1997
		ZA 9502980 A	11-01-1996
WO 9408036 A	14-04-1994	NONE	
EP 0513884 A	19-11-1992	AU 663871 B	26-10-1995
		AU 1698992 A	17-11-1992
		BR 9205894 A	27-09-1994
		CA 2105592 A	17-10-1992
		CZ 9302145 A	13-07-1994
		HU 65482 A	28-06-1994
		JP 6506595 T	28-07-1994
		WO 9218625 A	29-10-1992
		SK 111693 A	02-02-1994
WO 9421794 A	29-09-1994	AU 687793 B	05-03-1998
		AU 6262394 A	11-10-1994
		CA 2158473 A	29-09-1994
		EP 0689594 A	03-01-1996
		JP 8507923 T	27-08-1996
WO 9008828 A	09-08-1990	AU 1628695 A	03-08-1995
		AU 655574 B	05-01-1995
		AU 5037290 A	24-08-1990
		EP 0456706 A	21-11-1991
		JP 9182589 A	15-07-1997
		JP 4504355 T	06-08-1992
		US 5728558 A	17-03-1998
		US 5728926 A	17-03-1998
		US 5741684 A	21-04-1998
		US 5356799 A	18-10-1994
WO 9723599 A	03-07-1997	AU 1423997 A	17-07-1997
WO 9315599 A	19-08-1993	AU 3602593 A	03-09-1993

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 3,17,48,49 encoding cinnamate 4-hydroxylase (C4H) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

2. Claims: 1-54 partially

Isolated DNA sequences of ID nos 18,50-52 encoding coumarate 3-hydroxylase (C3H) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs, and transgenic plants and plant cells containing them.

3. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 35,36,81 encoding phenolase (PNL) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants, transgenic plants and plant cells containing said constructs.

4. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 22-25,53-55 encoding O-methyl transferase (OMT),plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

5. Claims: 1-54 all partially

Isolated DNA sequence of ID no 30, encoding cinnamyl alcohol dehydrogenase (CAD),plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

6. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 26-29,58-70 encoding

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NZ 97/00112

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☒ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

12. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 13,42-44,85-88 encoding peroxidase (POX) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

13. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 19-21 encoding ferulate-5-hydroxylase (F5H),plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

cinnamoyl-CoA reductase (CCR) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

7. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 16,45-47 encoding phenylalanine ammonia lyase (PAL),plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

8. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 56,57 encoding 4-coumarate:CoA ligase (4CL),plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

9. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 31-33,72 encoding coniferol glucosyl transferase (CGT),plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

10. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 34,73-80 encoding coniferin beta-glucosidase (CBG) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.

11. Claims: 1-54 all partially

Isolated DNA sequences of ID nos 37-41,82-84 encoding laccase (LAC) ,plant expression constructs incorporating said sequences, methods to modulate lignin content, structure, and enzyme activity in plants using said constructs,and transgenic plants and plant cells containing them.